Optical Interferometric Polarimetry I: Foundation

Nicholas M. Elias II

Astrometry Department, Navy Prototype Optical Interferometer

United States Naval Observatory

P.O. Box 1149

Flagstaff, AZ 86002-1149

Email: nme@nofs.navy.mil

Accepted by the Astrophysical Journal

ABSTRACT

The theoretical response of long-baseline optical interferometers and Fourier-transform spectrometers, including polarization effects, is derived. The formalism, employing the Jones and Mueller calculi as well as the relationship between them, was adapted from previous work in radio interferometry and tailored specifically for optical wavelengths. Expressions for Jones and Mueller matrices corresponding to specific optical components and effects are stated. It was determined that the system squared visibility depends on the instrument, atmosphere, and normalized intrinsic polarization of sources under observation. A sample algebraic calculation was performed to highlight the typical functional form for the instrumental system squared visibility, demonstrating that reductions from unity can be determined directly from differential measurements of polarization quantities between the arms. Monte-Carlo simulations were performed using two trains of identical mirrors with random relative-orientation offsets, yielding results consistent with the algebraic example. Four mathematical appendices are provided for reference.

Subject headings: techniques: interferometric — techniques: polarimetric

This series of papers is dedicated to the late Dr. William Blitzstein, astronomer and friend.

1. Introduction

Coherence is a fundamental property of electromagnetic radiation, providing a wealth of astrophysical information unavailable using incoherent techniques. Long-baseline interferometers, Fourier-transform spectrometers, and polarimeters are designed to observe different types of coherence. Long-baseline interferometers measure spatial coherence, in order reconstruct the structure of a source. Fourier-transform spectrometers measure temporal coherence, so that the frequency dependence of a source may be determined. Polarimeters measure the relative amounts of polarized and unpolarized flux, which is equivalent to measuring the degree of temporal coherence between orthogonal polarization states.

Polarization plays a role in interferometry at all wavelengths. For modern radio interferometers, electromagnetic radiation is incident upon an array of antennas (linear electric-field detectors), each consisting of one or more reflectors, polarized feeds, and waveguides. The radiation intercepted by each antenna is converted to an electrical signal, digitized, and multiplicatively correlated (interfered), with gain factors and phase shifts introduced at each step along the way. Michelson (phase-sensitive) optical interferometers do not convert visible light into electrical signals before combination, because the signal-to-noise ratio of the fringes will always be less than unity due to the Uncertainty Principle (Heffner 1962; Oliver 1965). Therefore, visible light must be interfered optomechanically (additively), not electronically, in such instruments. For modern optical interferometers, e.g., the Navy Prototype Optical Interferometer (NPOI; Armstrong *et al.* 1998), radiation is incident upon light collectors and reflected along a series of mirrors to one or more points of beam combination, finally ending at a square-law electric-field detector. Note that the polarizing and retarding properties of the components in optical interferometers are analogous to the gains and phase shifts in

radio interferometers (in the former case, however, the gains and phase shifts occur mostly before detection).

Over the past two decades, several groups have investigated polarization in relation to optical interferometers. Vakili (1983) described a polarimeter used in conjunction with the CERGA instruments. Also, Traub (1988), Beckers (1990), Rousselet-Perraut *et al.* (1996), and Ridgway and Bagnuolo (1996) derived simple expressions for polarization effects on fringe contrast. At radio wavelengths, Hamaker *et al.* (1996), Sault *et al.* (1996), and Hamaker and Bregman (1996) created an extensive theoretical foundation for understanding polarization in interferometry, including response and calibration. Some aspects of these three papers are applicable to optical interferometers.

For a greater understanding of instrumental capabilities, improved accuracy in astrophysical results, and the eventual ability to measure the spatial structure of a source in polarized visible-light (optical interferometric polarimetry, hereafter OIP), a theoretical framework specifically tailored for optical interferometers is both desirable and necessary. In this paper (Paper I), I will describe the total response of an optical interferometer in terms of the Jones and Mueller calculi (Shurcliff 1962; Kliger, Lewis, and Randall 1990) and perform example computer simulations using this formalism. The concepts and formulae in Paper I will be used extensively in the other papers of this series.

2. Spatial Coherence in Terms of Mueller Calculus

In "classical" optical interferometry theory (Armstrong *et al.* 1998, and references contained therein), the vector properties of electric fields are ignored when describing the struc-

ture of a source and the response of an instrument. This approximation is valid as long as the source is not highly polarized and the physical characteristics of each arm, including polarization and retardation effects, are not too dissimilar.

Let $E(\vec{\rho}_1, t; \kappa)$ and $E(\vec{\rho}_2, t; \kappa)$ represent the complex non-vector electric fields incident on each arm of a basic two-element optical interferometer (Figure 1), where $\vec{\rho}_1$ and $\vec{\rho}_2$ are the light-collector location vectors in a plane perpendicular to the source line of sight and $\kappa = 1/\lambda$ is the wavenumber. As these electric fields pass through their respective atmospheric paths and arms, they are modified until beam combination and detection; let these modifications be expressed as the complex multiplicative factors $J_1(\kappa)$ and $J_2(\kappa)$. The total electric field is the sum of the individual electric fields, or

$$E(\vec{\rho}_{12}, t; \kappa) = J_1(\kappa)\, E(\vec{\rho}_1, t; \kappa) + J_2(\kappa)\, E(\vec{\rho}_2, t; \kappa)\, e^{j2\pi\,\kappa\,\delta z(t)}\,, \tag{1}$$

where $j = \sqrt{-1}$, $\vec{\rho}_{12} = \vec{\rho}_1 - \vec{\rho}_2$, and $\delta z(t)$ is the delay-line dither across the central fringe (the delay line, included in $J_2(\kappa)$, has already removed the geometric delay).

Optical detectors are sensitive only to flux (the time average of the squared electric field), so the response of a basic two-element optical interferometer, integrating over the bandpass, is

$$I(\vec{\rho}_{12}, \delta z(t); \kappa_{eff}) = [M_{11}(\kappa_{eff}) + M_{22}(\kappa_{eff})]\, I(\vec{\mathbf{0}}; \kappa_{eff}) \;+\; 2\, M_{12}(\kappa_{eff})\, I(\vec{\rho}_{12}; \kappa_{eff})\, e^{j2\pi\,\kappa_{eff}\,\delta z(t)} \tag{2}$$

where $\kappa_{eff} = 1/\lambda_{eff}$ is the effective wavenumber of the bandpass, $I(\vec{\rho}_{12}, \delta z(t); \kappa_{eff})$ is the observed flux, $\vec{\mathbf{0}}$ is the null vector, $I(\vec{\mathbf{0}}; \kappa_{eff}) = |E(\vec{\rho}_1; \kappa_{eff})|^2 = |E(\vec{\rho}_2; \kappa_{eff})|^2$ is the flux measured by an ideal light collector (the arm or zero-spacing flux), $I(\vec{\rho}_{12}; \kappa_{eff}) = E(\vec{\rho}_1; \kappa_{eff})\, E^*(\vec{\rho}_2; \kappa_{eff})$ is the coherent flux measured by two ideal light collectors (the baseline flux), $M_{11}(\kappa_{eff}) = |J_1(\kappa_{eff})|^2$ and $M_{22}(\kappa_{eff}) = |J_2(\kappa_{eff})|^2$ are the arm throughputs, and

$M_{12}(\kappa_{eff}) = J_1(\kappa_{eff})\, J_2^*(\kappa_{eff})$ is the baseline throughput. In general, the throughputs may be functions of time. Note that $I(\vec{\rho}_{12}, \delta z(t); \kappa_{eff})$ is a real quantity, but I choose to express it as a phasor, for convenience. The bandpass, $\Delta\kappa = \kappa_{max} - \kappa_{min}$, is assumed small enough such that $I(\vec{\rho}_{12}; \kappa_{eff})$ is not reduced significantly (Tango and Twiss 1980). $\delta z(t)$ is used to distinguish between the arm and baseline terms during visibility measurements. It is varied over time scales $\tau_{int} \ll t_0$, where t_0 is the atmospheric temporal coherence (Buscher *et al.* 1995).

The observed visibility is defined as the ratio of the baseline and arm coefficients from Equation 2 (determined using Fourier techniques),

$$\begin{aligned} \mathcal{V}_{obs}(\vec{\rho}_{12}; \kappa_{eff}) &= \left[\frac{2\, M_{12}(\kappa_{eff})}{M_{11}(\kappa_{eff}) + M_{22}(\kappa_{eff})}\right] \left[\frac{I(\vec{\rho}_{12}; \kappa_{eff})}{I(\vec{\mathbf{0}}; \kappa_{eff})}\right] \\ &= \left[\frac{M_{12}(\kappa_{eff})}{M_0(\kappa_{eff})}\right] \left[\frac{I(\vec{\rho}_{12}; \kappa_{eff})}{I(\vec{\mathbf{0}}; \kappa_{eff})}\right] \\ &= K_{12}(\kappa_{eff})\, \mathcal{V}_s(\vec{\rho}_{12}; \kappa_{eff}), \end{aligned} \tag{3}$$

where $M_0(\kappa_{eff}) = \frac{1}{2}\,[M_{11}(\kappa_{eff}) + M_{22}(\kappa_{eff})]$ is the average arm throughput, $\mathcal{V}_s(\vec{\rho}_{12}; \kappa_{eff})$ $= I(\vec{\rho}_{12}; \kappa_{eff})/I(\vec{\mathbf{0}}; \kappa_{eff})$ is the source visibility, and $K_{12}(\kappa_{eff})$ is the system visibility or calibration constant ($0 \leq |K_{12}(\kappa_{eff})| \leq 1$). For ground-based optical interferometers, the phase of $\mathcal{V}_{obs}(\vec{\rho}_{12}; \kappa_{eff})$ fluctuates over time scales of order t_0 and $|\mathcal{V}_{obs}(\vec{\rho}_{12}; \kappa_{eff})|$ estimators include complicated bias corrections (Colavita 1985), so

$$|\mathcal{V}_{obs}(\vec{\rho}_{12}; \kappa_{eff})|^2 = |K_{12}(\kappa_{eff})|^2\, |\mathcal{V}_s(\vec{\rho}_{12}; \kappa_{eff})|^2 \tag{4}$$

is the principle observable quantity for modelling sources, especially for single-baseline configurations (assuming Poisson noise, it has a relatively simple bias correction; Colavita 1985). In practice, $|K_{12}(\kappa_{eff})|^2$ is significantly less than unity and time-dependent (due to low-frequency instrumental and atmospheric instability), so "calibration" sources (with a known

$|\mathcal{V}_s(\vec{\rho}_{12};\kappa_{eff})|^2$, preferably close to unity) are used to estimate $|K_{12}(\kappa_{eff})|^2$ for program-source observations.

The variables in Equations 1 and 2 are scalar approximations to vector quantities. To incorporate vectors into the previous analysis, Equation 1 must be expanded in terms of Jones vectors (cf. Appendix A) and Jones matrices (cf. Appendix B), or

$$\vec{\mathbf{E}}(\vec{\rho}_{12},t;\kappa) = \overleftrightarrow{\mathbf{J}}_1(\kappa)\cdot\vec{\mathbf{E}}(\vec{\rho}_1,t;\kappa) + \overleftrightarrow{\mathbf{J}}_2(\kappa)\cdot\vec{\mathbf{E}}(\vec{\rho}_2,t;\kappa)\,e^{j2\pi\,\kappa\,\delta z(t)}\,. \tag{5}$$

The scalar multiplication in Equation 2 may be replaced by a dot product, but substituting $\vec{\mathbf{E}}(\vec{\rho}_{12},t;\kappa) \rightarrow \vec{\mathbf{E}}'_A(t;\kappa)$ and $\vec{\mathbf{E}}^*(\vec{\rho}_{12},t;\kappa) \rightarrow \vec{\mathbf{E}}'^*_B(t;\kappa)$ (in Appendix B) and performing the necessary outer products and matrix transformations (also in Appendix B) yields a more general expression for the flux response of a basic optical interferometer,

$$\vec{\mathbf{S}}(\vec{\rho}_{12},\delta z(t);\kappa_{eff}) = \left[\overleftrightarrow{\mathbf{M}}_{11}(\kappa_{eff}) + \overleftrightarrow{\mathbf{M}}_{22}(\kappa_{eff})\right]\cdot\vec{\mathbf{S}}(\vec{\mathbf{0}};\kappa_{eff}) + 2\,\overleftrightarrow{\mathbf{M}}_{12}(\kappa_{eff})\cdot\vec{\mathbf{S}}(\vec{\rho}_{12};\kappa_{eff})\,e^{j2\pi\,\kappa_{eff}\,\delta z(t)}\,, \tag{6}$$

where the scalar fluxes have become 4×1 Stokes flux vectors and the scalar throughputs have become 4×4 Mueller throughput matrices. The polarization structure of the source is contained in this equation, which means that it is the starting point for OIP observations of astrophysically interesting objects. This topic will be treated in a later paper of this series.

The first element of $\vec{\mathbf{S}}(\vec{\rho}_{12},\delta z(t);\kappa_{eff})$, $I(\vec{\rho}_{12},\delta z(t);\kappa_{eff})$, is the quantity measured by detectors, so I define the detector operator $\hat{\mathbf{d}}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$ for the sake of mathematical convenience. The dot product of $\hat{\mathbf{d}}^T$ and Equation 6 (after factoring non-zero $I(\vec{\mathbf{0}};\kappa_{eff})$ and $I(\vec{\rho}_{12};\kappa_{eff})$ from their respective terms) results in Equation 2, so the scalar throughputs in Equation 3 may be equated to elements of the Mueller throughput matrices and normalized Stokes vectors of the source, specifically

$$
\begin{aligned}
M_{kl}(\kappa_{eff}) &= \hat{\mathbf{d}}^T \cdot \overleftrightarrow{\mathbf{M}}_{kl}(\kappa_{eff}) \cdot \vec{\mathbf{s}}(\vec{\rho}_{kl};\kappa_{eff}) \\
&= M_{kl}^{11}(\kappa_{eff}) + M_{kl}^{12}(\kappa_{eff})\, q(\vec{\rho}_{kl};\kappa_{eff}) + M_{kl}^{13}(\kappa_{eff})\, u(\vec{\rho}_{kl};\kappa_{eff}) + M_{kl}^{14}(\kappa_{eff})\, v(\vec{\rho}_{kl};\kappa_{eff})\,,
\end{aligned}
\tag{7}
$$

where the normalized Stokes vectors are defined in Appendix B and the superscripts are the row and column numbers, respectively, of the Mueller matrices. In other words, I have expressed the scalar version of the optical interferometer response in terms of the vector version, which will be useful for describing the effects of polarization on squared visibilities (classical non-OIP observations).

As an aside, I briefly discuss systematic errors introduced into $|K_{12}(\kappa_{eff})|^2$ due to source polarization. For the sake of argument, assume that the instrument and atmosphere are stable during calibrator and program source observations. If the sources have different polarizations, then $|K_{12}(\kappa_{eff})|^2$ estimated with the calibrator source will be slightly different from the correct value for the program source. Up to the present time, most sources observed with optical interferometers have polarizations less than 1%. Also, $M_{kl}^{12}(\kappa_{eff})$, $M_{kl}^{13}(\kappa_{eff})$, and $M_{kl}^{14}(\kappa_{eff})$ are $\sim$ one or two orders of magnitude smaller than $M_{kl}^{11}(\kappa_{eff})$. Therefore, systematics in $|K_{12}(\kappa_{eff})|^2$ due to source polarization are now of order 10^{-4} to 10^{-3}, which is nearly negligible when compared to typical observational errors. In the future, however, when sources with large (and variable!) polarizations are routinely observed, these systematic errors could become as large as 10^{-2}. *Therefore, accurate modelling of such program sources directly from squared visibilities may not be practical in the not-too-distant future, which means that standard interferometry imaging algorithms combined with OIP may become necessary for high-precision astrophysical results.*

The Mueller matrices for the arms, $\overleftrightarrow{\mathbf{M}}_{11}(\kappa_{eff})$ and $\overleftrightarrow{\mathbf{M}}_{22}(\kappa_{eff})$, may be expressed as the product of Mueller matrices corresponding to optical components (e.g., mirrors, lenses, etc.) and effects (e.g., the atmosphere). The baseline Mueller matrix, $\overleftrightarrow{\mathbf{M}}_{12}(\kappa_{eff})$, on the other hand, cannot be expressed in terms of simple Mueller matrices. Fortunately, there is an uncomplicated relationship between the Jones and Mueller calculi (cf. Appendix B), i.e., both the arm *and* baseline Mueller matrices may be expressed in terms of the Jones matrices of the arms, which themselves are products of Jones matrices for all optical components and effects. In Section 4, I will discuss explicit forms of individual Jones matrices relevant to optical interferometry, and in Section 6 I will employ the Jones matrices for mirrors in Monte-Carlo simulations.

3. Temporal Coherence in Terms of Mueller Calculus

A Fourier-transform spectrometer (FTS) is used to measure the wavenumber spectrum of a source (Mariotti and Ridgway 1988). Nordgren and Hajian (1999) configured the NPOI as an FTS with the addition of a few mirrors and a beam splitter. Specifically, light from a single siderostat was fed along both arms and interfered, and the intensity was measured as a function of delay-line position. If $\vec{\rho}_{12} = \vec{\mathbf{0}}$ and the wavenumber integration is shown explicitly, then Equation 2 becomes the response of an FTS,

$$\begin{aligned} I(D) &= \int_{\Delta\kappa} d\kappa \left[M_{11}(\kappa) + M_{22}(\kappa) + 2\, M_{12}(\kappa)\, e^{j2\pi\kappa D} \right] I(\kappa) \\ &= I_0 + 2 \int_{\Delta\kappa} d\kappa\, e^{j2\pi\kappa D}\, M_{12}(\kappa)\, I(\kappa)\,, \end{aligned} \tag{8}$$

where $\delta z(t) \rightarrow D$ is delay-line position, $I_0 = \int_{\Delta\kappa} d\kappa \left[M_{11}(\kappa) + M_{22}(\kappa) \right] I(\kappa)$ is an integration constant, $I(\kappa)$ is the source wavenumber spectrum, $M_{11}(\kappa)$ and $M_{22}(\kappa)$ are the arm through-

puts, and $M_{12}(\kappa)$ is the temporal-coherence throughput. Aside from a constant term and multiplicative factor, Equation 8 is just the Fourier transform and may be inverted,

$$I'(\kappa) = M_{12}(\kappa)\, I(\kappa) = \int_{\Delta D} dD\, e^{-j2\pi\kappa D}\, \frac{1}{2}\left[I(D) - I_0\right] , \tag{9}$$

where $I'(\kappa)$ is the observed wavenumber spectrum, $\Delta D \approx \frac{1}{2\Delta\kappa}$ is the delay range. I_0 is not known *a priori* because it depends on the source spectrum. For convenience, let $I_0 = \overline{I(D)}$ (the average), so that sampling artifacts from the zero-wavenumber component do not appear at other wavenumbers. If $M_{12}(\kappa) \approx$ *constant* over the bandpass, $I(\kappa)$ is known to within a multiplicative factor and the problem is finished, otherwise $M_{12}(\kappa)$ must be determined and its effects removed.

As in Section 2, Equation 9 may be generalized in terms of the Mueller calculus,

$$\vec{\mathbf{S}}'(\kappa) = \overleftrightarrow{\mathbf{M}}_{12}(\kappa) \cdot \vec{\mathbf{S}}(\kappa) = \int_{\Delta D} dD\, e^{-j2\pi\kappa D}\, \frac{1}{2}\left[\vec{\mathbf{S}}(D) - \vec{\mathbf{S}}_0\right] , \tag{10}$$

where the scalar fluxes have become 4×1 Stokes flux vectors and the scalar temporal-coherence throughput has become a 4×4 Mueller matrix (cf. Appendix B). Note that determining $\vec{\mathbf{S}}(\kappa)$ is more difficult than determining $I(\kappa)$ because a matrix inversion is involved. The dot product of $\hat{\mathbf{d}}^T$ and Equation 10 results in Equation 9, so $M_{12}(\kappa)$ may be expressed in terms of $\overleftrightarrow{\mathbf{M}}_{12}(\kappa)$ and the normalized Stokes vector of the source, namely

$$\begin{aligned} M_{12}(\kappa) &= \hat{\mathbf{d}}^T \cdot \overleftrightarrow{\mathbf{M}}_{12}(\kappa) \cdot \vec{\mathbf{s}}(\kappa) \\ &= M_{12}^{11}(\kappa) + M_{12}^{12}(\kappa)\, q(\kappa) + M_{12}^{13}(\kappa)\, u(\kappa) + M_{12}^{14}(\kappa)\, v(\kappa) . \end{aligned} \tag{11}$$

where the superscipts are the row and column numbers of the Mueller matrix. In a typical FTS, $M_{12}^{12}(\kappa)$, $M_{12}^{13}(\kappa)$, and $M_{12}^{14}(\kappa)$ are approximately one or two orders of magnitude smaller than $M_{12}^{11}(\kappa)$. If the source is only slightly polarized, less than 1%, then $M_{12}(\kappa) \approx M_{12}^{11}(\kappa)$.

On the other hand, if the source is highly polarized, then $M_{12}^{12}(\kappa)$, $M_{12}^{13}(\kappa)$, and $M_{12}^{14}(\kappa)$ must be determined with respect to $M_{12}^{11}(\kappa)$, otherwise the results will be significantly degraded. As in Section 2, $\overleftrightarrow{\mathbf{M}}_{12}(\kappa)$ may be expressed in terms of the total Jones matrices for each arm (cf. Appendix B).

4. Explicit Forms for Jones Matrices

The Jones matrix for the k^{th} arm of an optical interferometer, shown in Figure 1, may be expressed as the product of a number of Jones matrices,

$$\overleftrightarrow{\mathbf{J}}_k = \overleftrightarrow{\mathbf{J}}_{det,k} \cdot \overleftrightarrow{\mathbf{J}}_{bc,k} \cdot \overleftrightarrow{\mathbf{J}}_{train,k} \cdot \overleftrightarrow{\mathbf{J}}_{atm,k} \,, \tag{12}$$

where $\overleftrightarrow{\mathbf{J}}_{det,k}$ is due to the detector and the position of the beam on the detector, $\overleftrightarrow{\mathbf{J}}_{bc,k}$ is due to the beam-combiner input/output, $\overleftrightarrow{\mathbf{J}}_{train,k}$ is due to the train of mirror flats from the light collector to the last mirror before the beam combiner input/output, and $\overleftrightarrow{\mathbf{J}}_{atm,k}$ is due to the atmosphere above the light collector. Contained within these matrices are rotations due to basis transformations, which are discussed in detail in Appendix C. In Sections 4.1 and 4.2, I will describe these Jones matrices in their native $\{\hat{\mathbf{x}}\text{-}\hat{\mathbf{y}}\}$ bases.

4.1. Non-Mirrors

Assume that the atmosphere introduces a negligible amount of polarization and differential retardance upon an incoming wavefront. Therefore, the Jones matrix in its native basis is

$$\overleftrightarrow{\mathcal{J}}_{atm,k}(\kappa_{eff};t) = T_{atm,k}(\kappa_{eff};t)\; e^{j2\pi\,\bar{n}_k(\kappa_{eff};t)\,\kappa_{eff}\,L}\,\overleftrightarrow{\mathbf{1}} \,, \tag{13}$$

where $T_{atm,k}(\kappa_{eff};t)$ is the extinction integrated over the atmospheric path, $\bar{n}_k(\kappa_{eff};t)$ is the refractive index integrated over the atmospheric path, L is the atmospheric path length, and $\overleftrightarrow{\mathbf{1}}$ is the 2×2 unit matrix. $T_{atm,k}(\kappa_{eff};t)$ varies over time scales of seconds to hours, while $\bar{n}_k(\kappa_{eff};t)$ varies over time scales of order t_0.

The beam combiner may be modelled as a retarder plus weak polarizer, i.e.,

$$\overleftrightarrow{\mathcal{J}}_{bc,k}(\kappa_{eff}) = \begin{bmatrix} 1-\alpha_{bc,k}(\kappa_{eff}) & 0 \\ 0 & [1-\beta_{bc,k}(\kappa_{eff})]\, e^{j\psi_{bc,k}(\kappa_{eff})} \end{bmatrix}, \tag{14}$$

where $\psi_{bc,k}(\kappa_{eff})$ is the relative phase and $0 \leq \alpha_{bc,k}(\kappa_{eff})$, $\beta_{bc,k}(\kappa_{eff}) \leq 1$. Beam splitters are typically used as beam combiners, so each input/output combination (total of 4) will have different $\psi_{bc,k}(\kappa_{eff})$, $\alpha_{bc,k}(\kappa_{eff})$, and $\beta_{bc,k}(\kappa_{eff})$.

Some detectors exhibit intrinsic non-zero polarizibility, e.g., a solid-state detector with an anisotropic substrate. Also, non-zero polarizibility can be mimicked by beam motion on a detector with a position-dependent sensitivity. Therefore, I express the Jones matrix for a detector as

$$\overleftrightarrow{\mathcal{J}}_{det,k}(\kappa_{eff};x,y) = O_k(x,y) \begin{bmatrix} 1-\alpha_{det}(\kappa_{eff};x,y) & 0 \\ 0 & [1-\beta_{det}(\kappa_{eff};x,y)]\, e^{j\psi_{det}(\kappa_{eff};x,y)} \end{bmatrix}, \tag{15}$$

where x and y are the coordinates of the beam center on the detector, $\psi_{det}(\kappa_{eff};x,y)$ is the relative phase, $\alpha_{det}(\kappa_{eff};x,y)$ and $\beta_{det}(\kappa_{eff};x,y)$ are small positive constants, and $O_k(x,y)$ is the overlap function (assumed non-polarizing). $O_k(x,y)$ is defined such that $0 \leq O_k(x,y)\, O_l(x,y) \leq 1$ and $O_k^2(x,y) = 1$, and is used to model the effects of light dilution (squared-visibility reduction due to uninterfered light entering the detector). In real optical interferometers, x and y are functions of time because of angle-tracking errors.

Diffraction is a continuous process along the arms of an optical interferometer. For this reason, it was not included in Equation 12. If I assume that diffraction is non-polarizing, however, then the total effect may be modelled as as a single Jones matrix,

$$\overleftrightarrow{\mathcal{J}}_{diff,k}(\kappa_{eff}) = D_k(\kappa_{eff})\ \overleftrightarrow{\mathbf{1}}\,, \tag{16}$$

where $0 \leq |D_k(\kappa_{eff})| \leq 1$. Whenever a light beam enters/leaves a vacuum, it crosses a window. If I assume that the windows are non-polarizing, then their Jones matrices have the same form,

$$\overleftrightarrow{\mathcal{J}}_{win,km}(\kappa_{eff}) = W_{km}(\kappa_{eff})\ \overleftrightarrow{\mathbf{1}}\,, \tag{17}$$

where m is the window number.

4.2. Mirror Train

The Jones matrix for the mirror train, $\overleftrightarrow{\mathbf{J}}_{train,k}$, may be further subdivided into additional Jones matrices

$$\overleftrightarrow{\mathbf{J}}_{train,k} = \overleftrightarrow{\mathbf{J}}_{opt,k} \cdot \overleftrightarrow{\mathbf{J}}_{dl,k} \cdot \overleftrightarrow{\mathbf{J}}_{feed,k} \cdot \overleftrightarrow{\mathbf{J}}_{ao,k} \cdot \overleftrightarrow{\mathbf{J}}_{lc,k}\,, \tag{18}$$

where $\overleftrightarrow{\mathbf{J}}_{opt,k}$ is due to the mirrors that feed the beam combiner, $\overleftrightarrow{\mathbf{J}}_{dl,k}$ is due to the delay line, $\overleftrightarrow{\mathbf{J}}_{feed,k}$ is due to the feed system, $\overleftrightarrow{\mathbf{J}}_{ao,k}$ is due to the adaptive optics system, and $\overleftrightarrow{\mathbf{J}}_{lc,k}$ is due to the light collector. As before, all coordinate system transformations are contained within these Jones matrices (cf. Appendix C). The standard Jones matrix for a mirror in its native $\{\hat{\mathbf{x}}\text{-}\hat{\mathbf{y}}\} \rightarrow \{\hat{\mathbf{p}}\text{-}\hat{\mathbf{s}}\}$ basis may be found in Appendix D.

Delay lines typically consist of two mirrors (three reflections) in the same plane. Using the formulae in Appendix D, the Jones matrix of such a delay line, in its native basis, may

be expressed as

$$\begin{aligned} \overleftrightarrow{\mathcal{J}}_{dl,k}(\kappa_{eff};t) &\approx e^{j2\pi\,\kappa_{eff}\,[z_k(t)+\Delta z_k(t)]} \begin{bmatrix} r_p[i_1]\,r_p[i_2]\,r_p[i_1] & 0 \\ 0 & r_s[i_1]\,r_s[i_2]\,r_s[i_1] \end{bmatrix} \\ &= e^{j2\pi\,\kappa_{eff}\,[z_k(t)+\Delta z_k(t)]}\,\overleftrightarrow{\mathcal{J}}'_{dl,k}(\kappa_{eff})\,, \end{aligned} \tag{19}$$

where r_p and r_s are the p and s transmission coefficients (independent variable $\tilde{n}(\kappa_{eff})$ not explicitly shown), i_1 and i_2 are the incidence/reflection angles, $z_k(t)$ is the correct delay-line position that removes the geometric delay above the atmosphere (I assume that the delay lines are *in vacuo*) and $\Delta z_k(t)$ is the $z_k(t)$ error signal.

Assume tip-tilt adaptive optics. Tip-tilt mirrors move at most by a few arcseconds, which changes $\overleftrightarrow{\mathbf{J}}_{ao,k}$ by only a negligible amount. Of course, the mirrors cannot track the phases across the wavefronts perfectly, leading to a reduction in $|K_{12}(\kappa_{eff})|^2$. If the tip-tilt motions are approximately the same for both axes over time scales longer than t_0, then

$$\begin{aligned} \overleftrightarrow{\mathcal{J}}_{ao,k}(\kappa_{eff};t) &\approx A_k(t) \begin{bmatrix} r_p[i] & 0 \\ 0 & r_s[i] \end{bmatrix} \\ &= A_k(t)\,\overleftrightarrow{\mathcal{J}}'_{ao,k}(\kappa_{eff})\,, \end{aligned} \tag{20}$$

where i is the mean incidence angle and $A_k(t)$ is a complex function of time.

Light collectors may be either siderostats or telescopes. Up to the present time, most optical interferometers employ siderostats because of their well defined pivot points (leading to well defined baselines). For siderostats, the unit vectors defined in Appendix D still apply, except that they are functions of time that depend on the source coordinates and siderostat pointing model. If the orientation and/or polarization properties of the siderostat mirrors in a baseline are not the same, a small sky dependence in $|K_{12}(\kappa_{eff})|^2$ may result,

which introduces additional complications when calculating $|K_{12}(\kappa_{eff})|^2$. Both orientation and polarization problems may be virtually eliminated if the siderostats are replaced by telescopes with Cassegrain foci and the mirror trains replaced with polarization-preserving optical-fiber feed systems (Reynaud *et al.* 1994).

4.3. Multiplication and Factorization

In this section, functional dependencies will not be shown explicitly. I am not interested in effects due to the beam combiner and detector (cf. Section 6), so for simplicity I will assume that the beam combiner and detector are not polarizers/retarders, which means that $\gamma_{bc,k} = 1 - \alpha_{bc,k} = [1 - \beta_{bc,k}]\, e^{j\psi_{bc,k}}$ and $\gamma_{det,k} = 1 - \alpha_{det,k} = [1 - \beta_{det,k}]\, e^{j\psi_{det,k}}$. Therefore, the Jones matrix for the k^{th} arm of an optical interferometer is

$$\overset{\leftrightarrow}{\mathbf{J}}_k = D_k \left(\prod_m W_{km} \right) \gamma_{det,k}\; O_k\; \gamma_{bc,k}\; e^{j2\pi\,\kappa_{eff}\,[z_k + \Delta z_k]}\; A_k\; T_{atm,k}\; e^{j2\pi\,\kappa_{eff}\bar{n}_k L} \overset{\leftrightarrow}{\mathbf{J}}'_{train,k}\,, \qquad (21)$$

where

$$\overset{\leftrightarrow}{\mathbf{J}}'_{train,k} = \overset{\leftrightarrow}{\mathbf{J}}_{opt,k} \cdot \overset{\leftrightarrow}{\mathbf{J}}'_{dl,k} \cdot \overset{\leftrightarrow}{\mathbf{J}}_{feed,k} \cdot \overset{\leftrightarrow}{\mathbf{J}}'_{ao,k} \cdot \overset{\leftrightarrow}{\mathbf{J}}_{lc,k}\,. \qquad (22)$$

Using Equation 21 and the Jones-Mueller matrix relationship (cf. Appendix B), the scalar throughputs of Section 2 may be written as

$$M_{kl} = M_{diff,kl}\; M_{win,kl}\; M_{det,kl}\; M_{shear,kl}\; M_{bc,kl}\; M_{dl,kl}\; M_{ao,kl}\; M_{atm,kl}\; M'_{train,kl}\,, \qquad (23)$$

where $M_{diff,kl} = D_k\, D_l^*$, $M_{win,kl} = (\prod_m W_{km})\; (\prod_m W_{lm}^*)$, $M_{det,kl} = \gamma_{det,k}\, \gamma_{det,l}$, $M_{shear,kl} = O_k\, O_l$, $M_{bc,kl} = \gamma_{bc,k}\, \gamma_{bc,l}$, $M_{dl,kl} = e^{j2\pi\,\kappa_{eff}\,[z_{kl} + \Delta z_{kl}]}$, $M_{ao,kl} = A_k\, A_l^*$, $M_{atm,kl} = T_{atm,k}\, T_{atm,l}\, e^{j2\pi\,\kappa_{eff}\bar{n}_{kl} L}$, $M'_{train,kl} = \hat{\mathbf{d}}^T \cdot \overset{\leftrightarrow}{\mathbf{M}}'_{train,kl} \cdot \vec{\mathbf{s}}(\vec{\rho}_{kl})$, $\overset{\leftrightarrow}{\mathbf{M}}'_{train,kl} = \overset{\leftrightarrow}{\mathbf{T}} \cdot \left(\overset{\leftrightarrow}{\mathbf{J}}'_{train,k} \otimes \overset{\leftrightarrow}{\mathbf{J}}^{*\prime}_{train,l} \right) \cdot \overset{\leftrightarrow}{\mathbf{T}}^{-1}$, $\bar{n}_{kl} = \bar{n}_k - \bar{n}_l$, $z_{kl} = z_k - z_l$, and $\Delta z_{kl} = \Delta z_k - \Delta z_l$. Note that non-polarizing components/effects (Jones matrices

proportional to $\overleftrightarrow{\mathbf{1}}$) factor into their own scalar throughputs; Jones matrices for individual mirrors do not factor. If the scalar throughputs for each component/effect are similar for each arm, then

$$M_0 \approx M_{diff,0}\ M_{win,0}\ M_{det,0}\ M_{shear,0}\ M_{bc,0}\ M_{dl,0} M_{ao,0}\ M_{atm,0}\ M'_{train,0}\,, \tag{24}$$

where I define $M_{diff,0} = \frac{1}{2}\,[M_{diff,11} + M_{diff,22}]$, $\Delta M_{diff} = [M_{diff,11} - M_{diff,22}]$, etc. All of the first-order terms cancel, and I have neglected the higher-order terms. Using Equation 3, the calibration constant may then be written as

$$K_{12} \approx K_{diff,12}\ K_{win,12}\ K_{det,12}\ K_{shear,12}\ K_{bc,12}\ K_{dl,12}\ K_{ao,12}\ K_{atm,12}\ K'_{train,12}\,, \tag{25}$$

where $K_{diff,12} = M_{diff,12}/M_{diff,0}$, etc.

5. Simple Analytic Example

With the formalisms developed in Section 2 and Appendix B, I can derive a closed-form $|K_{12}|^2$ for an idealized optical interferometer. Let $\vec{\mathbf{E}}$ be an unpolarized electric field incident on both light collectors. For simplicity, assume that the Jones matrix of arm 1 is diagonal,

$$\overleftrightarrow{\mathbf{J}}_1 = \begin{bmatrix} r_a & 0 \\ 0 & r_b\, e^{j\phi} \end{bmatrix}\,, \tag{26}$$

where $0 \leq r_a,\ r_b \leq 1$ and ϕ is the relative phase. The transmission coefficients r_a and r_b are not directly measureable, so I express them in terms of throughputs, i.e., $R_a = \frac{1}{2}\,r_a^2$, $R_b = \frac{1}{2}\,\left|r_b\, e^{j\phi}\right|^2 = \frac{1}{2}\,r_b^2$, and $R = R_a + R_b$. The linear polarizability of arm 1 is defined as

$$p = \frac{R_b - R_a}{R_b + R_a}\,, \tag{27}$$

which is the amount of linear polarization produced when unpolarized light is incident upon the arm. With these variables, $\overleftrightarrow{\mathbf{J}}_1$ may be rewritten as

$$\overleftrightarrow{\mathbf{J}}_1(R,p,\phi) = \sqrt{R}\begin{bmatrix} \sqrt{1-p} & 0 \\ 0 & \sqrt{1+p}\,e^{j\phi} \end{bmatrix}. \tag{28}$$

The output electric field of arm 1 is $\vec{\mathbf{E}}_1 = \overleftrightarrow{\mathbf{J}}_1 \cdot \vec{\mathbf{E}}$.

Next, define a Jones matrix for arm 2, $\overleftrightarrow{\mathbf{J}}_2' = \overleftrightarrow{\mathbf{J}}_1(R+\Delta R, p+\Delta p, \phi+\Delta\phi)$. Just before combination, assume that the beam of arm 2 is rotated by a small angle ($\Delta\theta$) with respect to the beam of arm 1. Therefore, $\vec{\mathbf{E}}_2' = \overleftrightarrow{\mathbf{J}}_2' \cdot \vec{\mathbf{E}}$ and $\vec{\mathbf{E}}_2 = \overleftrightarrow{\mathbf{R}}(\Delta\theta) \cdot \vec{\mathbf{E}}_2' = \overleftrightarrow{\mathbf{R}}(\Delta\theta) \cdot \overleftrightarrow{\mathbf{J}}_2' \cdot \vec{\mathbf{E}} = \overleftrightarrow{\mathbf{J}}_2 \cdot \vec{\mathbf{E}}$, where $\overleftrightarrow{\mathbf{R}}(\Delta\theta)$ is a rotation matrix (cf. Appendix C).

Now that $\overleftrightarrow{\mathbf{J}}_1$ and $\overleftrightarrow{\mathbf{J}}_2$ have been specified, I calculate the exact form of $|K_{12}|^2$. For an unpolarized source, the throughputs in Equation 7 depend only on the M_{kl}^{11} terms of the Mueller matrices. Equation 3 then becomes

$$|K_{12}|^2 = \frac{1}{4}\,\frac{1+\frac{\Delta R}{R}}{\left[1+\frac{1}{2}\frac{\Delta R}{R}\right]^2}\left|(1-p)\sqrt{1-\frac{\Delta p}{1-p}} + (1+p)\sqrt{1+\frac{\Delta p}{1+p}}\,e^{-j\Delta\phi}\right|^2 \cos^2\Delta\theta\,. \tag{29}$$

For small p, $\Delta R/R$, $\Delta p/(1\pm p)$, $\Delta\phi$, and $\Delta\theta$, a simplified expression may be derived, namely

$$|K_{12}|^2 = 1 \;-\; \frac{1}{4}\left[\frac{\Delta R}{R}\right]^2 \;-\; \frac{1}{4}\Delta p^2 \;-\; \frac{1}{4}\Delta\phi^2 \;-\; \Delta\theta^2\,. \tag{30}$$

Note that all first-order terms cancel and any differences between the arms reduce $|K_{12}|^2$, as expected. Although this equation is useful for understanding general trends, keep in mind that $|K_{12}|^2$ for other system configurations will have slightly different forms. These results lead to an important concept, namely that *differential throughput and elliptical-polarization measurements can be used to predict* $|K_{12}|^2$ *for properly modelled* $\overleftrightarrow{\mathbf{J}}_1$ *and* $\overleftrightarrow{\mathbf{J}}_2$.

6. Monte-Carlo Mirror-Train Simulations

In this section, I perform example simulations dealing with the polarization and retardation effects of mirror trains in optical interferometers, further demonstrating the formalisms developed in Section 2 and Appendices B, C, and D. The simulation program, written in C++, is available by request from the author.

6.1. Description

The simplest optical interferometer has two near-identical mirror trains (cf. Figure 1), so for the purpose of these simulations I allowed the locations of the mirrors in the second train to be identical to the first except for a constant offset (cf. Figure 2). Since this paper does not deal with any specific optical interferometer, I performed a large number of Monte-Carlo runs using "random" mirror trains. A random mirror train consists of a series of N mirrors separated by identical distances and at random directions in three-dimensional space (cf. Figure 2); for these simulations, $2 \leq N \leq 15$. Approximately 10000 runs were required for each N to provide meaningful statistics.

The distribution of inclination angles for randomly distributed mirrors is shown in Figure 3 (the histogram symmetric about 45°). This distribution is not entirely acceptable, since typical optical interferometers do not have many (if any) large inclination angles. To remedy this situation, I artifically eliminated inclination angles greater than 60° and empirically skewed the distributions to slightly lower inclination angles (the other histogram, Figure 3).

Since each optical interferometer employs different types of mirror surfaces and operates over a variety of wavelengths (including the near-IR), I decided that the best compromise

for my simulations was uncoated silver in green light ($\tilde{n} = 0.177 - j3.638$ for $\kappa_{eff} = 1.697$ μm^{-1}; Jenkins and White 1957). Typical optical interferometers use mirrors with superior polarization and retardance properties, so these results represent worst-case scenarios. As a simplification, I assumed that all of the mirrors were pristine and identical, since I could not specify how the reflecting surfaces degraded with time. Others have discussed the effects of non-identical reflecting surfaces. Rousselet-Perraut *et al.* (1996) calculated that $\left|K'_{train,12}\right|^2$ (cf. Equation 25) is reduced by $\approx 1\%$ if the differential throughput between the arms is $\approx$ 20%, a relatively small effect (and consistent with Equation 30). Also, Beckers (1990) determined that differential phases between mirror trains have a significant effect on $\left|K'_{train,12}\right|^2$ (again consistent with Equation 30).

Because the mirrors have identical reflecting surfaces, the only other free parameters are the azimuth and elevation angle differences between corresponding mirrors in the arms. These deviations in orientation produce cumulative polarization and retardance differences between the trains, affecting $\left|K'_{train,12}\right|^2$ significantly (cf. Section 6.2). Here is a description of how the angle offsets for each run were applied. First, I created the two mirror trains. Second, I "twiddled" the first mirror of train 2 in both azimuth and elevation by deviations selected from Gaussian distributions whose 1σ widths (σ_{angle}) were between 0.5° and 3.0°. Third, I compensated for this misalignment by adjusting the next mirror to achieve parallelism. Normally, two following mirrors are used to achieve both parallelism and overlap, but for this analysis I assume that beam shear is not a polarizing effect, so I neglect it here (it factors out; cf. Section 4.3). Next, I twiddle the second mirror, and recursively repeat the process along the entire train. Note that the last mirror of train 2 was used only to force its output beam parallel to the output beam of train 1, which is all that is required because I am not considering the effects of beam combination.

6.2. Results

In this section, "distribution" bars are shown in all plots, i.e., the standard deviations about the means, not the errors of the means. In Figure 4, I show $M'_{train,11}$ versus N. Note that the distribution bars are very small (on the order of a few mmag) and the extinction is a linear function of the number of mirrors (slope $\approx$ 55 mmag mirror^{-1}). These results prove that mirror orientations have very little effect on $M'_{train,11}$, so these numbers may be used with little error for any optical interferometer with the same N and mirror surfaces. Figures 5 and 6 represent the linear polarization ($p_l = \sqrt{q^2 + u^2}$) and total polarization ($p_t = \sqrt{q^2 + u^2 + v^2}$), respectively, versus N for train 1. Note that after $N \gtrsim 10$, the polarizations first level off and then decrease. This situation is analogous to a system consisting of a series of weak polarizers at random orientations about the optical axis, i.e., as the number of components increases, the more the polarization effects cancel. Note that there are no plots for linear-polarization position angle (θ) and circular polarization (v), because their averages are undefined and 0.0, respectively, for random mirror trains.

In Figure 7, I plot $\Delta M'_{train} = M'_{train,22} - M'_{train,11}$, the RMS differential extinction between the two arms, versus N and σ_{angle}. Even for the worst case, $\Delta M'_{train} < 100$ μmag. Figures 8, 9, 10, and 11 display RMS differences of the linear polarization (Δp_l), linear-polarization position angle ($\Delta\theta$), total polarization (Δp_t), and circular polarization (Δv), respectively, between the mirror trains as a function of N and σ_{angle}.

In Figures 12 and 13, I show $\left|K'_{train,12}\right|^2$ versus N for $\sigma_{angle} = 1.0°$ and 3.0°, respectively. Note that the means decrease linearly with increasing N, and that the means decrease more quickly for $\sigma_{angle} = 3.0°$, as expected. Besides the means, the quantities of most interest are the distribution bar sizes and number of simulations above and below the means.

These quantities indicate that large reductions in $\left|K'_{train,12}\right|^2$ are possible and occur for a significant number of random feed systems. Large decreases in $\left|K'_{train,12}\right|^2$ are produced by large differential field rotations, amplifying polarization and retardance differences between corresponding mirrors (Beckers 1990). In Figure 14, I plot the mean $\left|K'_{train,12}\right|^2$ versus N over a range of σ_{angle} between 0.5° and 3.0° (no distribution bars, for clarity).

Because of the monotonic behavior of the mean $\left|K'_{train,12}\right|^2$ and the RMS differences, it should be possible to express the mean $\left|K'_{train,12}\right|^2$ in terms of these RMS differences, independent of N and σ_{angle}. Figures 15, 16, 17, 18, and 19 clearly show overlap to better than 1% for all N and σ_{angle}. These results lead to the same conclusion as stated in Section 5, namely that differential throughput and elliptical-polarization measurements may be used to predict the system squared visibility.

The polarization properties of mirrors plus Equation 30 can be used to derive an order-of-magnitude estimate of $\left|K'_{train,12}\right|^2$ versus σ_{angle}, a useful relationship for optical interferometer design. For a single mirror, $M'_{train,11}$, p_l, p_t, and v are $\propto i$ (the incidence angle) over the range of interest (between 0° and 60° degrees). I expect the same linear behavior for the differential quantities, i.e. $|\delta M'_{train}|$, $|\delta p_l|$, $|\delta p_t|$, and $|\delta v|$ versus $|\delta i|$, between a single pair of corresponding mirrors. In addition, $|\delta\theta| \propto |\delta i|$ and $|\delta\psi|$ (the differential rotation of the $\hat{\mathbf{p}}$ and $\hat{\mathbf{s}}$ unit vectors between a pair of mirrors; cf. Appendix D). Both $|\delta i|$ and $|\delta\psi| \propto \sigma_{angle}$, but what about the RMS Δi and $\Delta\psi$ for an entire mirror train? If I assume a two-dimensional "random-walk" model, then RMS $\Delta i \propto \sqrt{2N}\ |\delta i|$ and RMS $\Delta\psi \propto \sqrt{2N}\ |\delta\psi|$, which means that both RMS Δi and $\delta\psi$ are $\propto \sqrt{2N}\,\sigma_{angle}$. Because RMS $\Delta M'_{train}$, Δp_l, Δp_t, $\Delta\theta$, and Δv for the mirror train $\propto$ RMS Δi and $\Delta\psi$, they are also $\propto \sqrt{N}\,\sigma_{angle}$. Let $\Delta M'^2_{train} \rightarrow \Delta R^2$, $\Delta p_l^2 \rightarrow \Delta p^2$, and Δp_t^2, $\Delta v^2 \propto \Delta\phi^2$. When these quantities (and $\Delta\theta^2$) are substituted into

Equation 30, they give

$$\left|K'_{train,12}\right|^2 \approx 1 - fN\sigma^2_{angle}\,, \tag{31}$$

where f is a factor dependent on the type of mirror coating and the deviation from the random walk model. Note that this equation is consistent with Figure 14. Even if I used an optical system in Section 5 that led to a slightly different version of Equation 30, the functional form of Equation 31 would not change, which means that it is valid for any mirror train.

7. Conclusions

In this paper, a formalism was developed for modelling the response of an optical interferometer and Fourier-transform spectrometer, including polarization and retardation effects. The formalism includes both the Jones and Mueller calculi and the relationship between them. System squared visibilities depend not only on the instrument and atmosphere, but on the intrinsic polarization of sources as well. A sample algebraic calculation of the system squared visibility was performed, and a typical functional form for small differences between the arms was derived. These results showed that differential polarization measurements could become a useful tool for improving the accuracy of squared-visibility measurements for any specific instrument. Example Monte-Carlo simulations were performed for random mirror trains. These simulations, consistent with the algebraic calculation, showed that 1) significant reductions in $\left|K'_{train,12}\right|^2$ are possible for a significant number of random configurations and 2) σ_{angle} is an important design parameter for optical interferometers.

Once a polarization/retardation model for a specific optical interferometer is constructed, how can it be used? First, consider a covariance analysis, i.e., changing each mirror orien-

tation angle slightly in arm 2 to see how $\left|K'_{train,12}\right|^2$ is affected. Of course, after a mirror is altered, the next mirror in the train should be adjusted for parallelism. Second, in Section 4.2, I touched on the possibility of a sky-dependent $\left|K'_{train,12}\right|^2$. If polarization measurements of unpolarized sources were performed with each siderostat (not simultaneously), changes in $\left|K'_{train,12}\right|^2$ could be mapped across the sky and automatically applied to observed squared visibilities. This technique for measuring $\left|K'_{train,12}\right|^2$ has distinct advantage over measuring calibration-star visibilities, namely that it is immune to atmospheric effects.

According to Figure 7, the RMS $\Delta M'_{train}$ is extremely small, so it is not practical to measure it. The RMS quantities Δp_l, Δp_t, and Δv are not quite so small, and measureable with some effort. The last RMS quantity, $\Delta\theta$, on the other hand, is very easy to measure, so it should have slightly more weight when determining $\left|K'_{train,12}\right|^2$. Effects other than mirror orientation deviations, such as dust layers and coating degradation, could be included in the mirror-train analysis, if desired.

The author thanks Dr. Tyler E. Nordgren for drawing Figures 1 and 2; Drs. H. Melvin Dyck, Marvin E. Germain, Robert H. Koch, David Mozurkewich, Tyler E. Nordgren, Thomas A. Pauls, and Jeffrey J. Sudol, for critical review of this manuscript; and Gregory A. Shelton for obtaining important journal articles via interlibrary loan.

Appendices

A. Electric-Field/Jones-Vector Definitions

In general, starlight is partially polarized and may be represented as a sum of unpolarized and polarized electric fields,

$$\vec{\mathbf{E}}(\vec{\mathbf{r}},t) = \vec{\mathcal{E}}(\vec{\mathbf{r}},t) + \vec{\epsilon}(\vec{\mathbf{r}},t), \tag{A1}$$

where $\vec{\mathbf{E}}(\vec{\mathbf{r}},t)$ is the partially polarized electric field, $\vec{\mathcal{E}}(\vec{\mathbf{r}},t)$ is the unpolarized electric field, $\vec{\epsilon}(\vec{\mathbf{r}},t)$ is the polarized electric field, $\vec{\mathbf{r}} = [x,y,z]$ is the light-collector location, and t is the time. According to most elementary textbooks (e.g., Hecht 1987; Kliger, Lewis, and Randall 1990), only monochromatic plane waves (100% polarized) may be expressed as Jones vectors. If we are careful, however, Jones vectors may be extended to represent partially polarized and unpolarized plane-wave packets as well,

$$\begin{bmatrix} E_1(\vec{\mathbf{r}},t;\kappa) \\ E_2(\vec{\mathbf{r}},t;\kappa) \end{bmatrix} = \begin{bmatrix} \mathcal{E}_1(\vec{\mathbf{r}},t;\kappa) \\ \mathcal{E}_2(\vec{\mathbf{r}},t;\kappa) \end{bmatrix} + \begin{bmatrix} \epsilon_1(\vec{\mathbf{r}},t;\kappa) \\ \epsilon_2(\vec{\mathbf{r}},t;\kappa) \end{bmatrix}, \tag{A2}$$

where κ is the center wavenumber of the plane-wave packet and the subscripts indicate components in an orthogonal basis $\{\hat{\mathbf{1}}\text{-}\hat{\mathbf{2}}\}$ (e.g., horizontal/vertical, right-circular/left-circular). As an aside, I mention that Beckers (1990) uses the sum of specific 100%-polarized Jones vectors to mimic unpolarized light (when the Stokes vector of the sum is calculated, it is unpolarized).

Unpolarized light (more appropriately called natural light) is produced by a very large number of independent atomic emitters. Each excited atom radiates a randomly oriented polarized plane-wave packet at approximately $\kappa = \Delta E/hc$, where ΔE is the difference between energy levels, h is Planck's constant, and c is the speed of light. Each individual emission

lasts for roughly $\tau_e \sim 10^{-8}$ s, which determines the natural line width of the plane-wave packet $d\kappa/\kappa \sim 1/c\tau_e\kappa$ ($\sim 10^{-6}$ at optical wavelengths). All emissions with nearly the same wavenumber combine to form net polarized plane-wave packets which persist for no more than a few times τ_e. If the integration time $\tau_{int} \gg \tau_e$, the polarization information is lost, effectively yielding unpolarized light. Because $d\kappa$ is so small, it may be used as an integration variable over a bandpass width $\Delta\kappa$ (approximately centered on a bandpass effective wave number κ_{eff}).

An unpolarized electric field vector exhibits these characteristics: 1) components in any orthogonal basis produce the same time-integrated flux; and 2) the relative phase between components in any orthogonal basis varies randomly and rapidly with time, i.e., the temporal coherence is so low that for $\tau_{int} \gg \tau_e$ the time-integrated coherence is 0. Therefore, the Jones vector for unpolarized light may be represented as

$$\begin{bmatrix} \mathcal{E}_1(\vec{\mathbf{r}}, t; \kappa) \\ \mathcal{E}_2(\vec{\mathbf{r}}, t; \kappa) \end{bmatrix} = \begin{bmatrix} \mathcal{E}(\vec{\rho}; \kappa)\, e^{j\phi_1(\vec{\rho}, t; \kappa)} \\ \mathcal{E}(\vec{\rho}; \kappa)\, e^{j\phi_2(\vec{\rho}, t; \kappa)} \end{bmatrix} e^{\pm j2\pi\,(\kappa z - \nu t)}\,, \tag{A3}$$

where $\mathcal{E}(\vec{\rho}; \kappa)$ is the real amplitude for both components, the $\phi_i(\vec{\rho}, t; \kappa)$ are the random and temporally uncorrelated phases for the components, ν is the frequency, and $\vec{\rho} = [x, y]$. For convenience, I write the electric-field components as phasors. The $\{\hat{\mathbf{1}}\text{-}\hat{\mathbf{2}}\}$ basis and $\vec{\rho}$ lie in the $\hat{\mathbf{x}}$-$\hat{\mathbf{y}}$ plane and the wave propagates along the $\hat{\mathbf{z}}$ axis. The plane-wave packet width $d\kappa$ depends solely on the form of phases $\phi_i(\vec{\rho}, t; \kappa)$. The choice of $\pm$ is arbitrary, but consistency is essential (cf. Appendices B, C, and D).

Polarized light is produced when 1) atoms emit a significant number of photons with the same polarization; or 2) light interacts with matter. Unlike unpolarized light, orthogonal components of a polarized electric field are fully coherent, i.e., the relative phase between the

orthogonal components is constant. Therefore, the Jones vector of a polarized plane-wave packet may be expressed in phasor notation as

$$\vec{\epsilon}(\vec{\mathbf{r}}, t; \kappa) = \begin{bmatrix} \epsilon_1(\vec{\rho}; \kappa) \\ \epsilon_2(\vec{\rho}; \kappa)\, e^{j\phi(\vec{\rho};\kappa)} \end{bmatrix} e^{\pm j2\pi\,(\kappa z - \nu t)}\,, \tag{A4}$$

where $\epsilon_1(\vec{\rho}; \kappa)$ and $\epsilon_2(\vec{\rho}; \kappa)$ are the real electric-field component amplitudes and $\phi(\vec{\rho}; \kappa)$ is the constant phase difference. The width of this plane-wave packet is arbitrary, so I choose $d\kappa$ (the same as before). Note that unpolarized and polarized electric fields are temporally incoherent with each other, by definition. This fact is used in the time averages in Appendix B.

B. Jones, Coherence, and Stokes Relationships

Consider a Jones vector (an electric-field wave packet written in phasor notation; cf. Appendix A) that has been modified by an optical system, or

$$\vec{\mathbf{E}}'(t; \kappa) = \overleftrightarrow{\mathbf{J}}(\kappa) \cdot \vec{\mathbf{E}}(t; \kappa)\,, \tag{B1}$$

where

$$\overleftrightarrow{\mathbf{J}}(\kappa) = \begin{bmatrix} J^{11}(\kappa) & J^{12}(\kappa) \\ J^{21}(\kappa) & J^{22}(\kappa) \end{bmatrix} \tag{B2}$$

is the complex Jones matrix of the optical system in the $\hat{\mathbf{1}}$-$\hat{\mathbf{2}}$ basis. For more details concerning the Jones calculus, see Shurcliff (1962) and Kliger, Lewis, and Randall (1990). With these equations and the definitions of Appendix A, the coherence of two arbitrary electric fields A and B, integrated over time and wavenumber bandpass, may be expressed as (Hamaker *et*

al. 1996)

$$
\begin{aligned}
\vec{\mathbf{C}}'(\kappa_{eff}) &= \int_{\Delta\kappa} d\kappa \left\langle \vec{\mathbf{E}}'_A(t;\kappa) \otimes \vec{\mathbf{E}}'^*_B(t;\kappa) \right\rangle \\
&= \int_{\Delta\kappa} d\kappa \left\langle \left[\overleftrightarrow{\mathbf{J}}_A(\kappa) \cdot \vec{\mathbf{E}}_A(t;\kappa) \right] \otimes \left[\overleftrightarrow{\mathbf{J}}^*_B(\kappa) \cdot \vec{\mathbf{E}}^*_B(t;\kappa) \right] \right\rangle \\
&= \int_{\Delta\kappa} d\kappa \left[\overleftrightarrow{\mathbf{J}}_A(\kappa) \otimes \overleftrightarrow{\mathbf{J}}^*_B(\kappa) \right] \cdot \left\langle \vec{\mathbf{E}}_A(t;\kappa) \otimes \vec{\mathbf{E}}^*_B(t;\kappa) \right\rangle \\
&= \left[\overleftrightarrow{\mathbf{J}}_A(\kappa_{eff}) \otimes \overleftrightarrow{\mathbf{J}}^*_B(\kappa_{eff}) \right] \cdot \left\langle \vec{\mathbf{E}}_A(t;\kappa_{eff}) \otimes \vec{\mathbf{E}}^*_B(t;\kappa_{eff}) \right\rangle \\
&= \overleftrightarrow{\mathcal{C}}(\kappa_{eff}) \cdot \begin{bmatrix} \mathcal{E}_A(\kappa_{eff})\mathcal{E}_B(\kappa_{eff}) + \epsilon_{A1}(\kappa_{eff})\epsilon_{B1}(\kappa_{eff}) \\ \epsilon_{A1}(\kappa_{eff})\epsilon_{B2}(\kappa_{eff})\, e^{-j\phi_B(\kappa_{eff})} \\ \epsilon_{A2}(\kappa_{eff})\epsilon_{B1}(\kappa_{eff})\, e^{j\phi_A(\kappa_{eff})} \\ \mathcal{E}_A(\kappa_{eff})\mathcal{E}_B(\kappa_{eff}) + \epsilon_{A2}(\kappa_{eff})\epsilon_{B2}(\kappa_{eff})\, e^{j\left[\phi_A(\kappa_{eff})-\phi_B(\kappa_{eff})\right]} \end{bmatrix} \\
&= \overleftrightarrow{\mathcal{C}}(\kappa_{eff}) \cdot \begin{bmatrix} \frac{1}{2}\mathcal{I}(\kappa_{eff}) + i_{11}(\kappa_{eff}) \\ i_{12}(\kappa_{eff}) \\ i_{21}(\kappa_{eff}) \\ \frac{1}{2}\mathcal{I}(\kappa_{eff}) + i_{22}(\kappa_{eff}) \end{bmatrix} \\
&= \overleftrightarrow{\mathcal{C}}(\kappa_{eff}) \cdot \vec{\mathbf{C}}(\kappa_{eff}) \,, \qquad \text{(B3)}
\end{aligned}
$$

where $\langle\rangle$ denotes the time average over τ_{int}, $\otimes$ is the outer (direct) product (Hamaker *et al.* 1996), $\overleftrightarrow{\mathcal{C}}(\kappa_{eff})$ is the coherence matrix, $\mathcal{I}(\kappa_{eff})$ is the unpolarized flux, and $i_{11}(\kappa_{eff})$, $i_{12}(\kappa_{eff})$, $i_{21}(\kappa_{eff})$, and $i_{22}(\kappa_{eff})$ are the polarized fluxes in the $\hat{\mathbf{1}}$-$\hat{\mathbf{2}}$ basis. In this derivation, I employed the theorem in Hamaker *et al.* (1996) to separate the Jones matrices from the Jones vectors.

The coherence vector may be converted into a Stokes vector by a simple matrix transformation,

$$
\vec{\mathbf{S}}'(\kappa_{eff}) = \overleftrightarrow{\mathbf{T}} \cdot \vec{\mathbf{C}}'(\kappa_{eff})
$$

$$
\begin{aligned}
&= \overleftrightarrow{\mathbf{T}} \cdot \overleftrightarrow{\mathcal{C}}(\kappa_{eff}) \cdot \vec{\mathbf{C}}(\kappa_{eff}) \\
&= \left[\overleftrightarrow{\mathbf{T}} \cdot \overleftrightarrow{\mathcal{C}}(\kappa_{eff}) \cdot \overleftrightarrow{\mathbf{T}}^{-1}\right] \cdot \left[\overleftrightarrow{\mathbf{T}} \cdot \vec{\mathbf{C}}(\kappa_{eff})\right] \\
&= \overleftrightarrow{\mathbf{M}}(\kappa_{eff}) \cdot \vec{\mathbf{S}}(\kappa_{eff}) \\
&= \overleftrightarrow{\mathbf{M}}(\kappa_{eff}) \cdot \begin{bmatrix} I(\kappa_{eff}) \\ Q(\kappa_{eff}) \\ U(\kappa_{eff}) \\ V(\kappa_{eff}) \end{bmatrix} \\
&= \begin{bmatrix} I'(\kappa_{eff}) \\ Q'(\kappa_{eff}) \\ U'(\kappa_{eff}) \\ V'(\kappa_{eff}) \end{bmatrix}
\end{aligned}
\tag{B4}
$$

where $\overleftrightarrow{\mathbf{M}}(\kappa_{eff})$ is the Mueller matrix; $I(\kappa_{eff})$, $Q(\kappa_{eff})$, $U(\kappa_{eff})$, and $V(\kappa_{eff})$ (and the corresponding primed quantities) are the Stokes parameter fluxes; and $\overleftrightarrow{\mathbf{T}}$ is the coherence-to-Stokes transformation matrix. Stokes vectors may also be written in normalized form,

$$
\begin{aligned}
\vec{\mathbf{s}}(\kappa_{eff}) &= \vec{\mathbf{S}}(\kappa_{eff}) \,/\, I(\kappa_{eff}) \\
&= \begin{bmatrix} 1 \\ q(\kappa_{eff}) \\ u(\kappa_{eff}) \\ v(\kappa_{eff}) \end{bmatrix},
\end{aligned}
\tag{B5}
$$

where $q(\kappa_{eff})$, $u(\kappa_{eff})$, and $v(\kappa_{eff})$ are the normalized Stokes parameter fluxes. For more details concerning the Mueller calculus, see Shurcliff (1962) and Kliger, Lewis, and Randall (1990).

The exact form of $\overset{\leftrightarrow}{\mathbf{T}}$ depends on the Jones vector basis. If $\hat{\mathbf{1}} \rightarrow \hat{\mathbf{x}}$ and $\hat{\mathbf{2}} \rightarrow \hat{\mathbf{y}}$, then

$$\overset{\leftrightarrow}{\mathbf{T}} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & \pm j & \mp j & 0 \end{bmatrix} \tag{B6}$$

and

$$\vec{\mathbf{S}}(\kappa_{eff}) = \begin{bmatrix} \mathcal{I}(\kappa_{eff}) + i_{xx}(\kappa_{eff}) + i_{yy}(\kappa_{eff}) \\ i_{xx}(\kappa_{eff}) - i_{yy}(\kappa_{eff}) \\ i_{xy}(\kappa_{eff}) + i_{yx}(\kappa_{eff}) \\ \pm j\left[i_{xy}(\kappa_{eff}) - i_{yx}(\kappa_{eff})\right] \end{bmatrix}. \tag{B7}$$

Alternatively, if $\hat{\mathbf{1}} \rightarrow \hat{\mathbf{R}} = \left[\hat{\mathbf{x}} \pm j\hat{\mathbf{y}}\right]/\sqrt{2}$ (right-circular polarization) and $\hat{\mathbf{2}} \rightarrow \hat{\mathbf{L}} = \left[\hat{\mathbf{x}} \mp j\hat{\mathbf{y}}\right]/\sqrt{2}$ (left-circular polarization), then

$$\overset{\leftrightarrow}{\mathbf{T}} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & \pm j & \mp j & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \tag{B8}$$

and

$$\vec{\mathbf{S}}(\kappa_{eff}) = \begin{bmatrix} \mathcal{I}(\kappa_{eff}) + i_{RR}(\kappa_{eff}) + i_{LL}(\kappa_{eff}) \\ i_{RL}(\kappa_{eff}) + i_{LR}(\kappa_{eff}) \\ \pm j\left[i_{RL}(\kappa_{eff}) - i_{LR}(\kappa_{eff})\right] \\ i_{RR}(\kappa_{eff}) - i_{LL}(\kappa_{eff}) \end{bmatrix}. \tag{B9}$$

Note that when $\vec{\mathbf{E}}_A(t;\kappa) = \vec{\mathbf{E}}_B(t;\kappa)$, Equations B7 and B9 become the classical (zero-spacing) Stokes vector definitions (Hecht 1987).

For the rest of this appendix, functional dependences will be suppressed for the sake of brevity. Seven elements of any Mueller matrix are most important for characterizing its effects, so I present them here. The most often quoted quantity is the element in the upper left-hand corner (row 1, column 1), corresponding to the throughput of unpolarized light,

$$M_{kl}^{11} = \frac{1}{2}\left[J_k^{11}J_l^{11*} + J_k^{12}J_l^{12*} + J_k^{21}J_l^{21*} + J_k^{22}J_l^{22*}\right] , \tag{B10}$$

where the superscripts are the row and column identifications, respectively. Note that this expression is independent of basis. The other three elements of the first row (from left to right), are the throughputs for the Stokes q, u, and v polarizations, respectively. In the $\{\hat{\mathbf{1}}\text{-}\hat{\mathbf{2}}\} \rightarrow \{\hat{\mathbf{x}}\text{-}\hat{\mathbf{y}}\}$ basis, these throughputs are given by,

$$M_{kl}^{12} = \frac{1}{2}\left[J_k^{xx}J_l^{xx*} - J_k^{xy}J_l^{xy*} + J_k^{yx}J_l^{yx*} - J_k^{yy}J_l^{yy*}\right] , \tag{B11a}$$

$$M_{kl}^{13} = \frac{1}{2}\left[J_k^{xx}J_l^{xy*} + J_k^{xy}J_l^{xx*} + J_k^{yx}J_l^{yy*} + J_k^{yy}J_l^{yx*}\right] , \tag{B11b}$$

and

$$M_{kl}^{14} = \mp j\,\frac{1}{2}\left[J_k^{xx}J_l^{xy*} - J_k^{xy}J_l^{xx*} + J_k^{yx}J_l^{yy*} - J_k^{yy}J_l^{yx*}\right] , \tag{B11c}$$

while in the $\{\hat{\mathbf{1}}\text{-}\hat{\mathbf{2}}\} \rightarrow \{\hat{\mathbf{R}}\text{-}\hat{\mathbf{L}}\}$ basis,

$$M_{kl}^{12} = \frac{1}{2}\left[J_k^{RR}J_l^{RL*} + J_k^{RL}J_l^{RR*} + J_k^{LR}J_l^{LL*} + J_k^{LL}J_l^{LR*}\right] , \tag{B12a}$$

$$M_{kl}^{13} = \mp j\,\frac{1}{2}\left[J_k^{RR}J_l^{RL*} - J_k^{RL}J_l^{RR*} + J_k^{LR}J_l^{LL*} - J_k^{LL}J_l^{LR*}\right] , \tag{B12b}$$

and

$$M_{kl}^{14} = \frac{1}{2}\left[J_k^{RR}J_l^{RR*} - J_k^{RL}J_l^{RL*} + J_k^{LR}J_l^{LR*} - J_k^{LL}J_l^{LL*}\right] . \tag{B12c}$$

The other three elements in the first column (from top to bottom), are the Stokes q, u, and v polarizibilities, respectively. The polarizibility is defined as the amount of polarization

produced when unpolarized light is incident on the system. In the $\{\hat{\mathbf{1}}\text{-}\hat{\mathbf{2}}\} \rightarrow \{\hat{\mathbf{x}}\text{-}\hat{\mathbf{y}}\}$ basis, these polarizibilities are given by,

$$M_{kl}^{21} = \frac{1}{2}\left[J_k^{xx} J_l^{xx*} + J_k^{xy} J_l^{xy*} - J_k^{yx} J_l^{yx*} - J_k^{yy} J_l^{yy*}\right] , \tag{B13a}$$

$$M_{kl}^{31} = \frac{1}{2}\left[J_k^{xx} J_l^{yx*} + J_k^{yx} J_l^{xx*} + J_k^{xy} J_l^{yy*} + J_k^{yy} J_l^{xy*}\right] , \tag{B13b}$$

and

$$M_{kl}^{41} = \pm j\,\frac{1}{2}\left[J_k^{xx} J_l^{yx*} - J_k^{yx} J_l^{xx*} + J_k^{xy} J_l^{yy*} - J_k^{yy} J_l^{xy*}\right] , \tag{B13c}$$

while in the $\{\hat{\mathbf{1}}\text{-}\hat{\mathbf{2}}\} \rightarrow \{\hat{\mathbf{R}}\text{-}\hat{\mathbf{L}}\}$ basis,

$$M_{kl}^{21} = \frac{1}{2}\left[J_k^{RR} J_l^{LR*} + J_k^{LR} J_l^{RR*} + J_k^{RL} J_l^{LL*} + J_k^{LL} J_l^{RL*}\right] , \tag{B14a}$$

$$M_{kl}^{31} = \pm j\,\frac{1}{2}\left[J_k^{RR} J_l^{LR*} - J_k^{LR} J_l^{RR*} + J_k^{RL} J_l^{LL*} - J_k^{LL} J_l^{RL*}\right] , \tag{B14b}$$

and

$$M_{kl}^{41} = \frac{1}{2}\left[J_k^{RR} J_l^{RR*} + J_k^{RL} J_l^{RL*} - J_k^{LR} J_l^{LR*} - J_k^{LL} J_l^{LL*}\right] . \tag{B14c}$$

Going from the Stokes/Mueller domain to the Jones domain is often useful, so I present a summary of the mathematics here. Each Mueller matrix element (M_{kl}^{mn}) may be written as an equation whose right-hand side consists of four terms (e.g., Equations B10, B11, B12, B13, and B14). Each of these four terms consists of a product of two Jones-matrix elements ($J_k^{pq} J_l^{rs}$). These 16 equations may be expressed in matrix form, i.e., a 16×1 column vector of M_{kl}^{mn} equal to a 16×16 block-diagonal matrix (each block 4×4) times a 16×1 column vector of $J_k^{pq} J_l^{rs}$. The solution for each block is simple, since they are unitary (their inverses are conjugate transposes). To solve for the individual Jones-matrix elements, first choose non-zero J_k^{mn} and J_l^{mn} (for the sake of argument, assume m, $n = 1$ here). Then for (m,n) = (1,2),

(2,1), and (2,2), calculate the ratios $R_k^{mn} = J_k^{mn}/J_k^{11}$ and $R_l^{mn} = J_l^{mn}/J_l^{11}$. Only three phases are required to uniquely describe a Jones matrix, so let J_k^{11} and J_l^{11} be real. For a zero-spacing Mueller matrix, $\overleftrightarrow{\mathbf{J}}_k = \overleftrightarrow{\mathbf{J}}_l \rightarrow J_k^{11} = J_l^{11}$, and the problem is solved ($J_k^{11}\, J_l^{11}$ is already known). For a baseline Mueller matrix, J_k^{11} and J_l^{11} cannot be determined independently unless the ratio of the unpolarized throughputs for each arm, M_{kk}^{11}/M_{ll}^{11}, is specified. Using Equation B10 for M_{kk}^{11} and M_{ll}^{11}, I conclude that

$$\frac{J_k^{11}}{J_l^{11}} = \sqrt{\frac{M_{kk}^{11}}{M_{ll}^{11}}}\,\sqrt{\frac{1+|R_l^{12}|^2+|R_l^{21}|^2+|R_l^{22}|^2}{1+|R_k^{12}|^2+|R_k^{21}|^2+|R_k^{22}|^2}}\,, \tag{B15}$$

which in conjunction with the product $J_k^{11}\, J_l^{11}$ completely determines $\overleftrightarrow{\mathbf{J}}_k$ and $\overleftrightarrow{\mathbf{J}}_l$.

C. Transformation and Rotation

Jones vectors and matrices may be transformed between bases using the matrix $\overleftrightarrow{\mathbf{H}}$,

$$\begin{aligned}
\vec{\mathbf{E}}'' &= \overleftrightarrow{\mathbf{H}} \cdot \overleftrightarrow{\mathbf{J}} \cdot \vec{\mathbf{E}} \\
&= \left[\overleftrightarrow{\mathbf{H}} \cdot \overleftrightarrow{\mathbf{J}} \cdot \overleftrightarrow{\mathbf{H}}^{-1}\right] \cdot \left[\overleftrightarrow{\mathbf{H}} \cdot \vec{\mathbf{E}}\right] \\
&= \overleftrightarrow{\mathbf{J}}' \cdot \vec{\mathbf{E}}'
\end{aligned} \tag{C1}$$

where $\vec{\mathbf{E}}''$ is the output Jones vector, $\vec{\mathbf{E}}$ is the initial Jones vector, $\vec{\mathbf{E}}'$ is the transformed Jones vector, $\overleftrightarrow{\mathbf{J}}$ is the Jones matrix, and $\overleftrightarrow{\mathbf{J}}'$ is the transformed Jones matrix. According to Hamaker and Bregman (1996), $\overleftrightarrow{\mathbf{H}}$ has a simple form when transforming from the $\{\hat{\mathbf{x}}\text{-}\hat{\mathbf{y}}\}$ basis to $\{\hat{\mathbf{R}}\text{-}\hat{\mathbf{L}}\}$ basis, namely

$$\overleftrightarrow{\mathbf{H}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & \mp j \\ 1 & \pm j \end{bmatrix}, \tag{C2}$$

where the signs depend on the sign of the propagating exponential in Appendix A. The inverse of this matrix, $\overleftrightarrow{\mathbf{H}}^{-1}$, does exist, and is equal to the conjugate transpose of $\overleftrightarrow{\mathbf{H}}$. I do

not employ a basis transformation matrix for the Mueller calculus (Hamaker *et al.* 1996) because I use different versions of the coherence-to-Stokes transformation matrix $\overleftrightarrow{\mathbf{T}}$ instead (both methods are equivalent; cf. Appendix B).

The coordinate system of Jones vectors and matrices may be rotated as well. Rotation matrices are analogous to basis transformation matrices, so $\overleftrightarrow{\mathbf{H}}$ in Equation C1 may be replaced by the rotation matrix $\overleftrightarrow{\mathbf{R}}(\psi)$, where ψ is the rotation angle. In the $\{\hat{\mathbf{x}}\text{-}\hat{\mathbf{y}}\}$ basis,

$$\overleftrightarrow{\mathbf{R}}(\psi) = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix}, \tag{C3}$$

while in the $\{\hat{\mathbf{R}},\hat{\mathbf{L}}\}$ basis

$$\overleftrightarrow{\mathbf{R}}(\psi) \rightarrow \overleftrightarrow{\mathbf{H}} \cdot \overleftrightarrow{\mathbf{R}}(\psi) \cdot \overleftrightarrow{\mathbf{H}}^{-1} = \begin{bmatrix} e^{\pm j\psi} & 0 \\ 0 & e^{\mp j\psi} \end{bmatrix}. \tag{C4}$$

The inverse of this matrix, $\overleftrightarrow{\mathbf{R}}(\psi)^{-1}$, is simply $\overleftrightarrow{\mathbf{R}}(-\psi)$. Note that $\overleftrightarrow{\mathbf{R}}(\psi)$ is a Jones matrix in its own right, namely a field rotator.

Stokes vector and Mueller matrix rotations are performed in exactly the same manner,

$$\begin{aligned} \vec{\mathbf{S}}'' &= \overleftrightarrow{\mathcal{R}} \cdot \overleftrightarrow{\mathbf{M}} \cdot \vec{\mathbf{S}} \\ &= \left[\overleftrightarrow{\mathcal{R}} \cdot \overleftrightarrow{\mathbf{M}} \cdot \overleftrightarrow{\mathcal{R}}^{-1}\right] \cdot \left[\overleftrightarrow{\mathcal{R}} \cdot \vec{\mathbf{S}}\right] \\ &= \overleftrightarrow{\mathbf{M}}' \cdot \vec{\mathbf{S}}'. \end{aligned} \tag{C5}$$

The rotation matrix is

$$\overleftrightarrow{\mathcal{R}}(\psi) = \overleftrightarrow{\mathbf{T}} \cdot \overleftrightarrow{\mathbf{R}}(\psi) \cdot \overleftrightarrow{\mathbf{T}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\psi & \sin 2\psi & 0 \\ 0 & -\sin 2\psi & \cos 2\psi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{C6}$$

The inverse transformation is $\overleftrightarrow{\mathcal{R}}(-\psi)$. This matrix is also a Mueller matrix representing a field rotator.

D. Mirrors and Mirror Trains

The incidence plane of a mirror, commonly called the p plane (p = parallel), contains the incident and reflected Poynting unit vectors $\hat{\kappa}^i$ and $\hat{\kappa}^r$, and the mirror-normal unit vector $\hat{\mathbf{m}}$. The s plane ($s = senkrecht$ = perpendicular, in German) is perpendicular to the p plane. In this paper, $\hat{\mathbf{p}} \rightarrow \hat{\mathbf{x}}$ and $\hat{\mathbf{s}} \rightarrow \hat{\mathbf{y}}$, so the Jones matrix for a mirror in its native basis is

$$\overleftrightarrow{\mathcal{J}}_{mirror}[i;\tilde{n}(\kappa_{eff})] = \begin{bmatrix} r_p[i;\tilde{n}(\kappa_{eff})] & 0 \\ 0 & r_s[i;\tilde{n}(\kappa_{eff})] \end{bmatrix}, \tag{D1}$$

where

$$r_p[i;\tilde{n}(\kappa_{eff})] = \frac{\tilde{n}^2(\kappa_{eff})\cos i - \sqrt{\tilde{n}^2(\kappa_{eff}) - \sin^2 i}}{\tilde{n}^2(\kappa_{eff})\cos i + \sqrt{\tilde{n}^2(\kappa_{eff}) - \sin^2 i}} \tag{D2a}$$

and

$$r_s[i;\tilde{n}(\kappa_{eff})] = \frac{\cos i - \sqrt{\tilde{n}^2(\kappa_{eff}) - \sin^2 i}}{\cos i + \sqrt{\tilde{n}^2(\kappa_{eff}) - \sin^2 i}} \tag{D2b}$$

are the Fresnel mirror coefficients, i is the angle of incidence ($i = 0°$ is normal incidence), and $\tilde{n}(\kappa_{eff}) = n_r(\kappa_{eff}) + jn_i(\kappa_{eff})$ is the mirror refractive index (Reitz *et al.* 1979). These formulae assume that the beam is incident from a vacuum to a thick (larger than the light wavelength) and homogeneous metal layer, which is acceptable here. For a perfect mirror (n_r and/or $n_i \rightarrow \infty$), $r_p \rightarrow 1$ and $r_s \rightarrow -1$ at normal incidence. Note that the signs of these coefficients depend on the directions of the corresponding unit vectors (cf. below).

$\tilde{n}(\kappa_{eff})$ may be expressed in terms of two measureable quantities,

$$\tilde{n}(\kappa_{eff}) = -\sin i_{pr}\sqrt{1 + \tan^2 i_{pr}\left[p(i_{pr};\kappa_{eff}) - j\sqrt{1 - p^2(i_{pr};\kappa_{eff})}\right]^2}, \tag{D3}$$

where i_{pr} is the principal incidence angle and $p(i_{pr};\kappa_{eff})$ is the principal polarization (the polarization produced when unpolarized light is incident at i_{pr}). When $i = i_{pr}$, $r_p[i_{pr};\tilde{n}(\kappa_{eff})]$ is at its minimum, $p(i_{pr};\kappa_{eff})$ is at its maximum, and the phase difference between $r_p[i_{pr};\tilde{n}(\kappa_{eff})]$ and $r_s[i_{pr};\tilde{n}(\kappa_{eff})]$ is exactly 90°. Note that most commercial mirrors employ a protective dielectric layer over the metal layer, which leads to effective $\tilde{n}(\kappa_{eff})$ that are somewhat different from textbook values of the underlying layers.

In general, mirrors in a train do not share the same coordinate system (basis). Assume that an electric field $\vec{\mathbf{E}}$ is incident upon a train of mirrors, and is expressed in terms of the basis of the first mirror. Then, the resultant electric field is

$$\begin{aligned} \vec{\mathbf{E}}' &= \overleftrightarrow{\mathbf{R}}(\psi_{N,N+1}) \cdot \overleftrightarrow{\mathbf{J}}_N \cdot \ \ldots \ \cdot \overleftrightarrow{\mathbf{R}}(\psi_{n-1,n}) \cdot \overleftrightarrow{\mathbf{J}}_{n-1} \cdot \ \ldots \ \cdot \overleftrightarrow{\mathbf{R}}(\psi_{1,2}) \cdot \overleftrightarrow{\mathbf{J}}_1 \cdot \vec{\mathbf{E}} \\ &= \overleftrightarrow{\mathbf{J}}' \cdot \vec{\mathbf{E}}, \end{aligned} \tag{D4}$$

where the $\overleftrightarrow{\mathbf{R}}()$ are rotation matrices found in Appendix C. The final rotation matrix, $\overleftrightarrow{\mathbf{R}}(\psi_{N,N+1})$, is used only for one arm of an optical interferometer baseline, and is required in order to express the $\vec{\mathbf{E}}'$ of one arm in terms of the coordinate system of the other arm.

The technique for determining the $\hat{\mathbf{p}}$ and $\hat{\mathbf{s}}$ rotations along a train of mirrors (Figure 2) is as follows. The *a priori* quantities are the first incident Poynting unit vector, $\hat{\kappa}_1^i$, and all of the mirror-normal unit vectors, $\hat{\mathbf{m}}_{n,n+1}$, where n is the mirror number. First, calculate the initial incidence (= reflectance) angle. Then, calculate the first reflected Poynting unit vector, $\hat{\kappa}_1^r$. Next, set $\hat{\kappa}_2^i = \hat{\kappa}_1^r$ and repeat this process for the entire mirror train. The incidence angle and reflected Poynting vector are given by

$$\cos i_n = -\,\hat{\kappa}_n^i \cdot \hat{m}_{n,n+1} \tag{D5}$$

and

$$\hat{\kappa}_n^r = \hat{\kappa}_{n+1}^i = \hat{\kappa}_n^i + (2\cos i_n)\ \hat{m}_{n,n+1}\,. \tag{D6}$$

Once these quantities have been determined, the incident s and p unit vectors, $\hat{\mathbf{s}}_n^i$ and $\hat{\mathbf{p}}_n^i$, must be calculated,

$$\hat{\mathbf{s}}_n^i = \frac{\hat{\kappa}_n^i \times \hat{\mathbf{m}}_{n,n+1}}{|\hat{\kappa}_n^i \times \hat{\mathbf{m}}_{n,n+1}|} \tag{D7}$$

and

$$\hat{\mathbf{p}}_n^i = -\,\hat{\kappa}_n^i \times \hat{\mathbf{s}}_n^i\,. \tag{D8}$$

There is a choice, on reflection, whether the s or p unit vectors remain unchanged. The standard convention is $\hat{\mathbf{s}}_n^r = \hat{\mathbf{s}}_n^i = \hat{\mathbf{s}}_n$, so

$$\hat{\mathbf{p}}_n^r = -\,\hat{\kappa}_n^r \times \hat{\mathbf{s}}_n\,. \tag{D9}$$

The rotation angles needed for Equation D4 are given by

$$\tan\psi_{n,n+1} = \frac{\hat{\kappa}_n^r \cdot \left[\hat{\mathbf{s}}_n \times \hat{\mathbf{s}}_{n+1}\right]}{\hat{\mathbf{s}}_n \cdot \hat{\mathbf{s}}_{n+1}}\,. \tag{D10}$$

Note than the "atan2" function should be used to calculate $\psi_{n,n+1}$, since it takes the quadrant into account.

REFERENCES

Armstrong, J.T., Mozurkewich, D., Rickard, L.J, Hutter, D.J., Benson, J.A., Bowers, P.F., Elias II, N.M., Hummel, C.A., Johnston, K.J., Buscher, D.F., Clark III, J.H., Ha, L., Ling, L.-C., White, N.M., and Simon, R.S. 1998, ApJ 496, 550.

Beckers, J. 1990, in *Polarization Considerations for Optical Systems II*, Proc. SPIE 1166, 380.

Buscher, D.F., Armstrong, J.T., Hummel, C.A., Quirrenbach, A., Mozurkewich, D., Johnston, K.J., Denison, C.S., Colavita, M.M., and Shao, M. 1995, AppOpt 34, 1081.

Colavita, M.M. 1985, *Ph.D. thesis*, Massachusetts Institute of Technology.

Hamaker, J.P., Bregman, J.D., and Sault, R.J. 1996, A&AS 117, 137.

Hamaker, J.P. and Bregman, J.D. 1996, A&AS 117, 161.

Hecht, E. 1987, *Optics*, 2nd edition (Reading, MA: Addison-Wesley).

Heffner, H. 1962, Proc. IRE 50, 1604.

Jenkins, F.A. and White, H.E. 1957, in *Fundamentals of Optics*, 3rd edition (New York: McGraw-Hill), pg. 524.

Kliger, D.S., Lewis, J.W., and Randall, C.E. 1990, *Polarized Light in Optics and Spectroscopy* (San Diego:Academic Press), pp. 59-152.

Mariotti, J.-M. and Ridgway, S.T. 1988, A&A 195, 350.

Nordgren, T.E. and Hajian, A.R. 1999, in *Precise Stellar Radial Velocities*, Proc. IAU Coll. 170, ed. J.B. Hearnshaw and C.D. (San Francisco:ASP), Vol. 185, pg. 36.

Oliver, B.M. 1965, Proc. IEEE 53, 436.

Reitz, J.R., Milford, F.J., Christy, R.W. 1979, in *Foundations of Electromagnetic Theory*, 3rd edition (Reading, MA: Addision-Wesley), pp. 382-401.

Reynaud, F., Alleman, J.J., and Lagorceix, H. 1994, Proc. of SPIE 2209, 431.

Ridgway, S.T. and Bagnuolo, W.G. 1996, CHARA Technical Report No. 28.

Roberts, D.H., Lehár, J., and Dreher, J.W. 1987, AJ 93, 968.

Rousselet-Perraut, K., Vakili, F., and Mourard, D. 1996, Opt.Eng. 35, 2943.

Sault, R.J., Hamaker, J.P., and Bregman, J.D. 1996, A&AS 117, 149.

Shurcliff, W.A. 1962, Polarized Light (Cambridge:Harvard University Press).

Tango, W.J. and Twiss, R.Q. 1980, in *Progress in Optics*, Vol. 17, ed. E. Wolf, p. 241.

Traub, W.A. 1988, in *High-Resolution Imaging by Interferometry*, ed. F. Merkle., Proc. ESO Conference and Workshop Proceedings No. 29, p. 1029.

Vakili, F. 1983, in *Instrumentation in Astronomy V*, eds. A, Boksenberg and D.L. Crawford, Proc. SPIE 445, p. 484.

This manuscript was prepared with the AAS LaTeX macros v5.0.

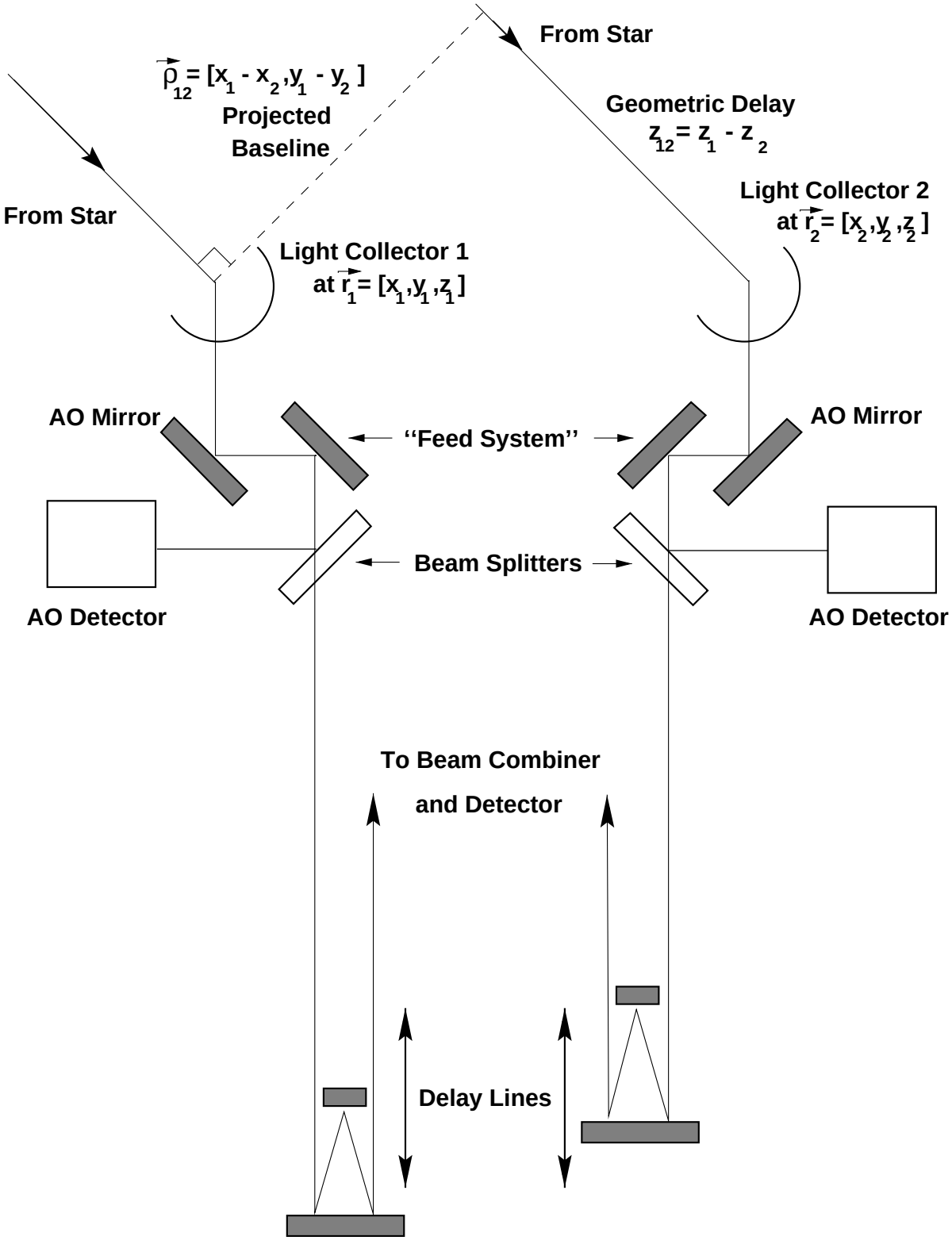

Fig. 1.— A schematic diagram of an ideal two-element optical interferometer.

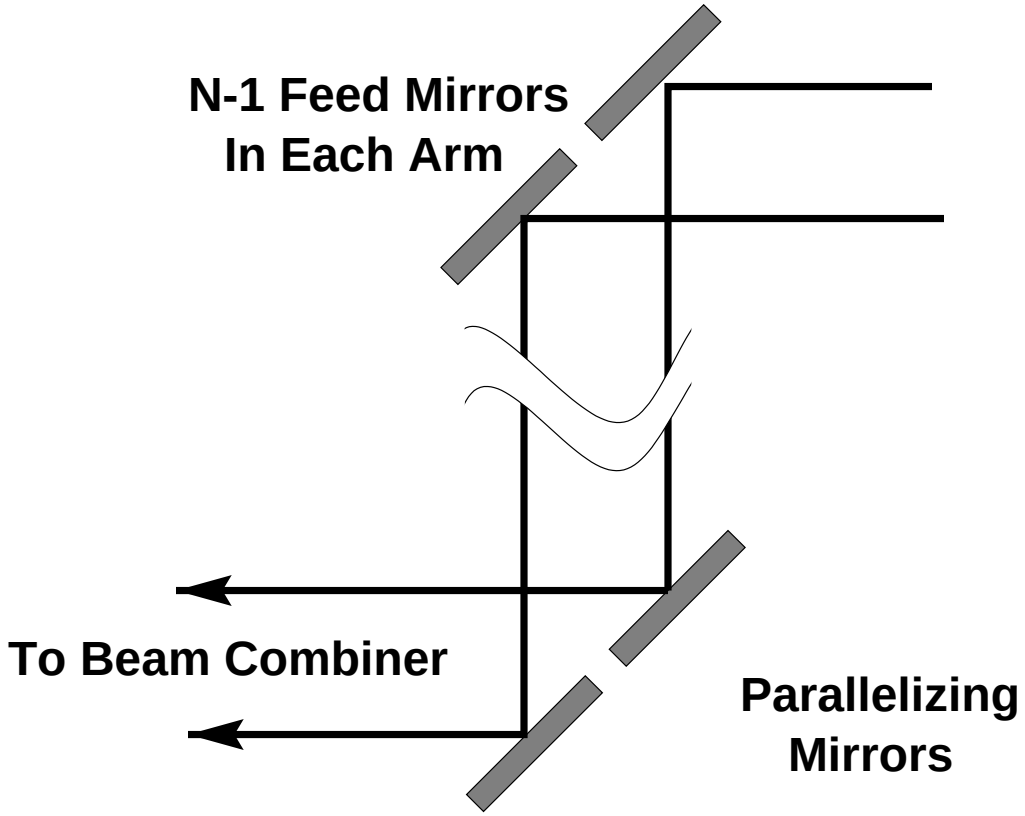

Fig. 2.— A schematic diagram of a random feed system for the simulations in this paper. The mirrors in arm 1 are separated by unit length, and their orientations are random. The mirrors in arm 2 are offset in space by a fixed amount from the corresponding mirrors in arm 1, plus they have small mirror orientation errors. The effects of a beam combiner are not considered.

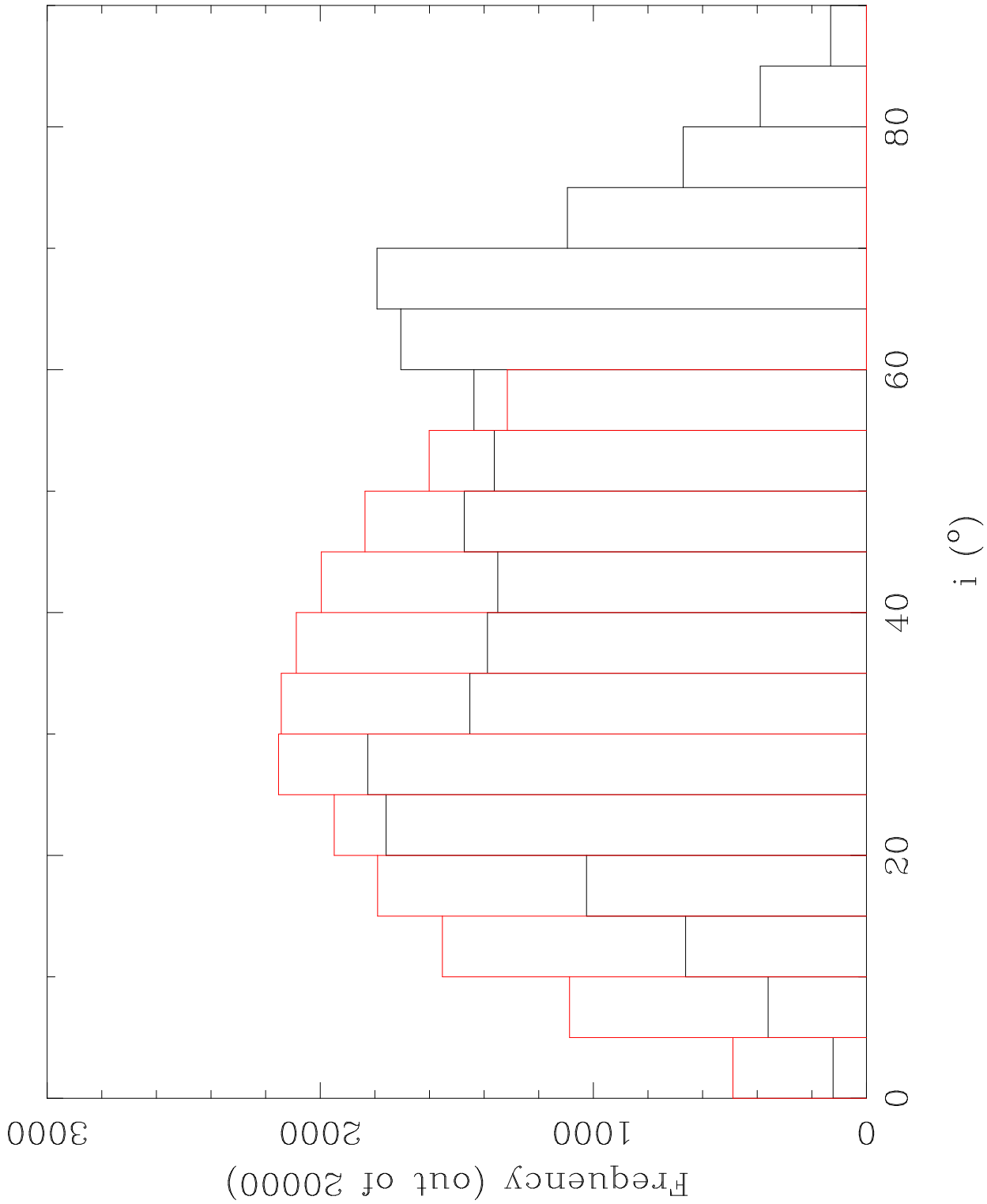

Fig. 3.— Incidence angle distributions, fully random (symmetric) and modified (asymmetric).

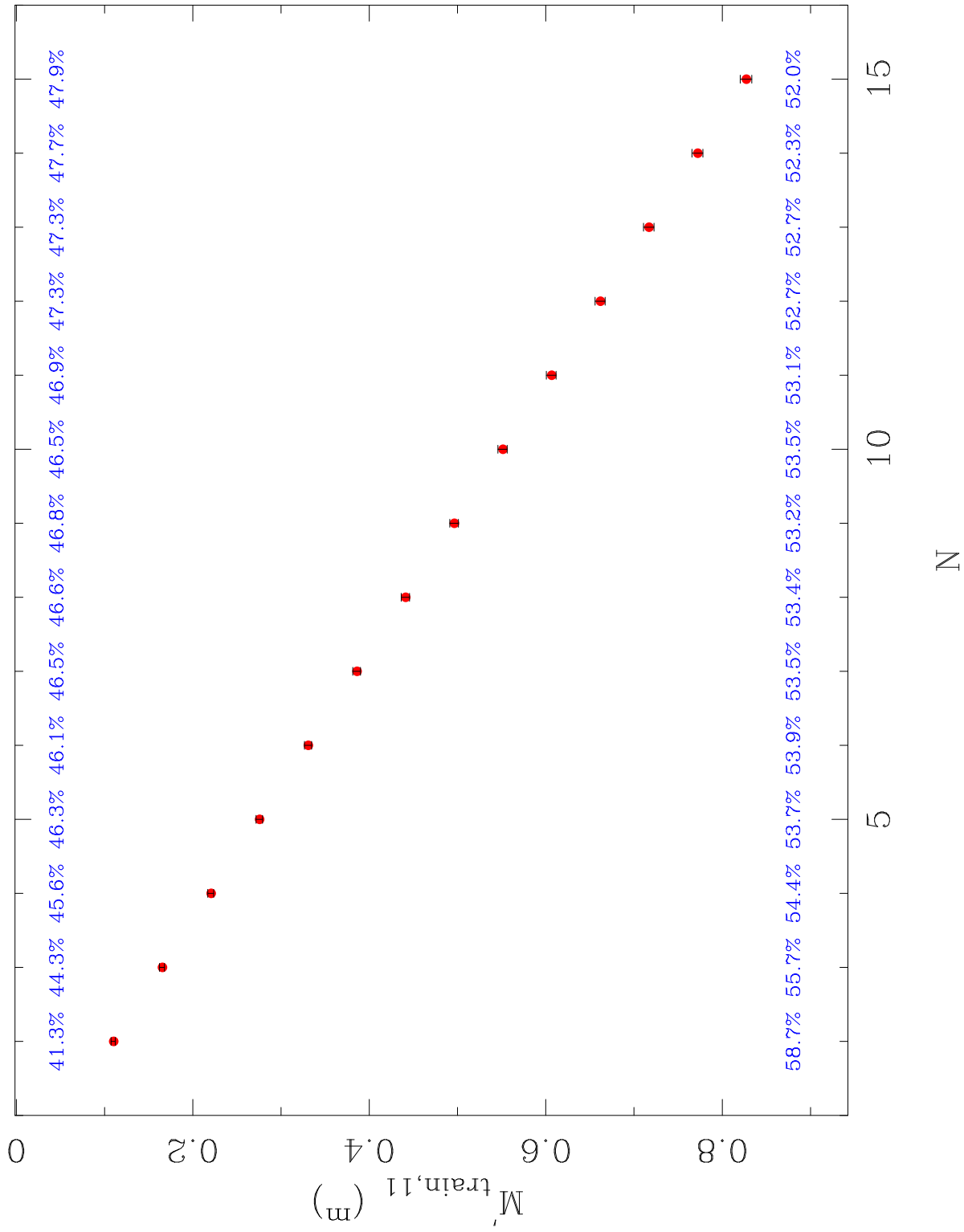

Fig. 4.— $M'_{train,11}$ versus N, arm 1. Distribution bars, not errors of the means, are shown. The percentages represent the probabilities above and below the mean.

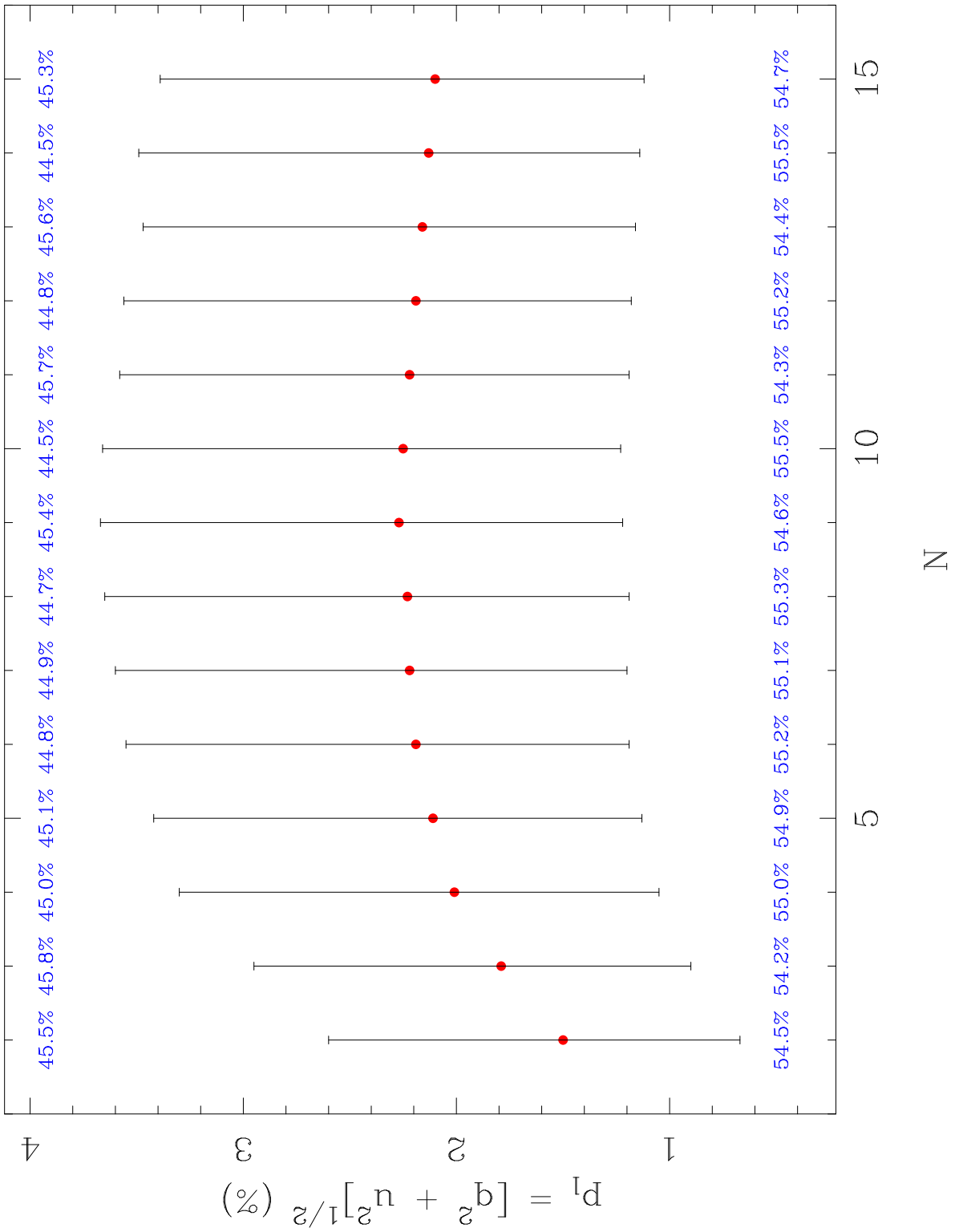

Fig. 5.— p_l versus N, arm 1. Distribution bars, not errors of the means, are shown. The percentages represent the probabilities above and below the mean.

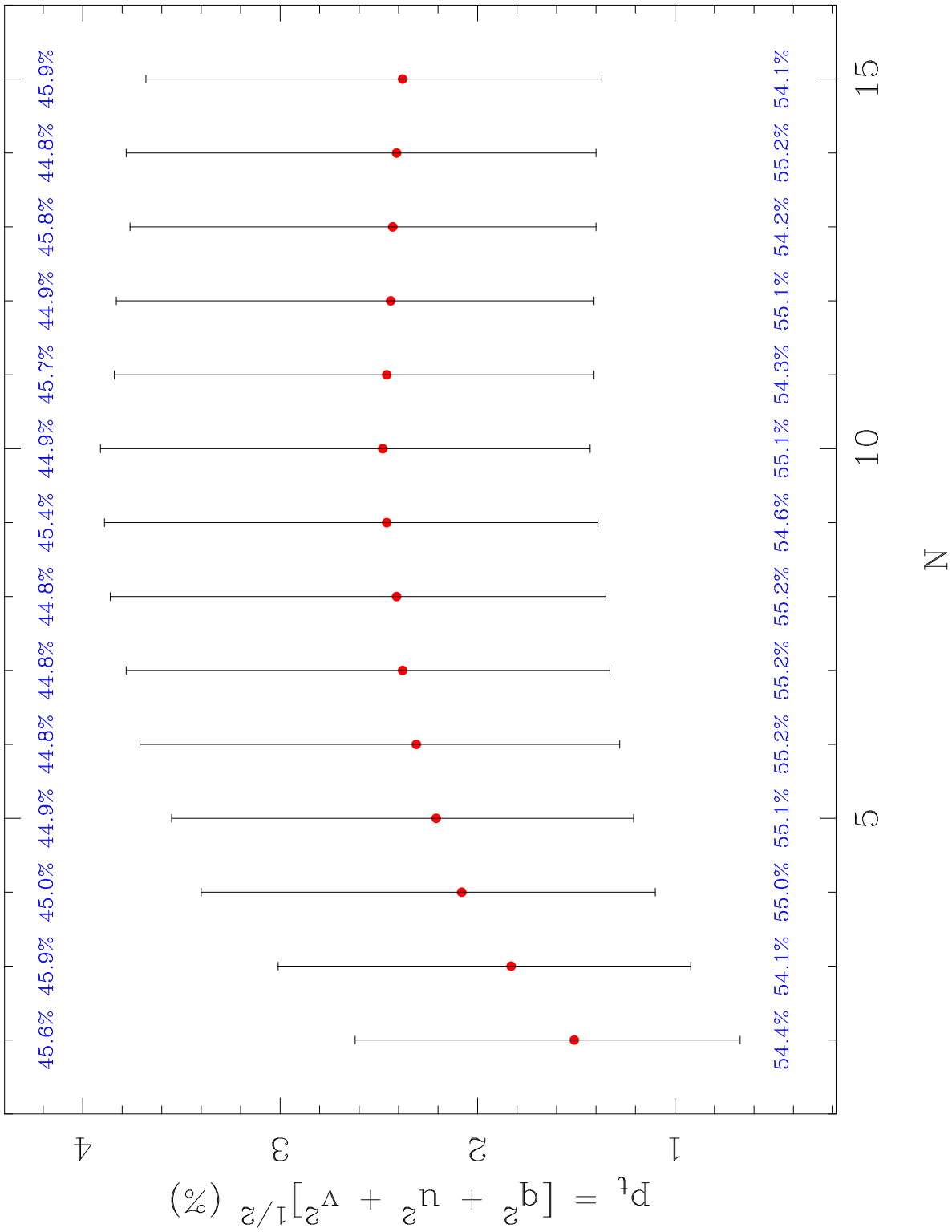

Fig. 6.— p_t versus N, arm 1. Distribution bars, not errors of the means, are shown. The percentages represent the probabilities above and below the mean.

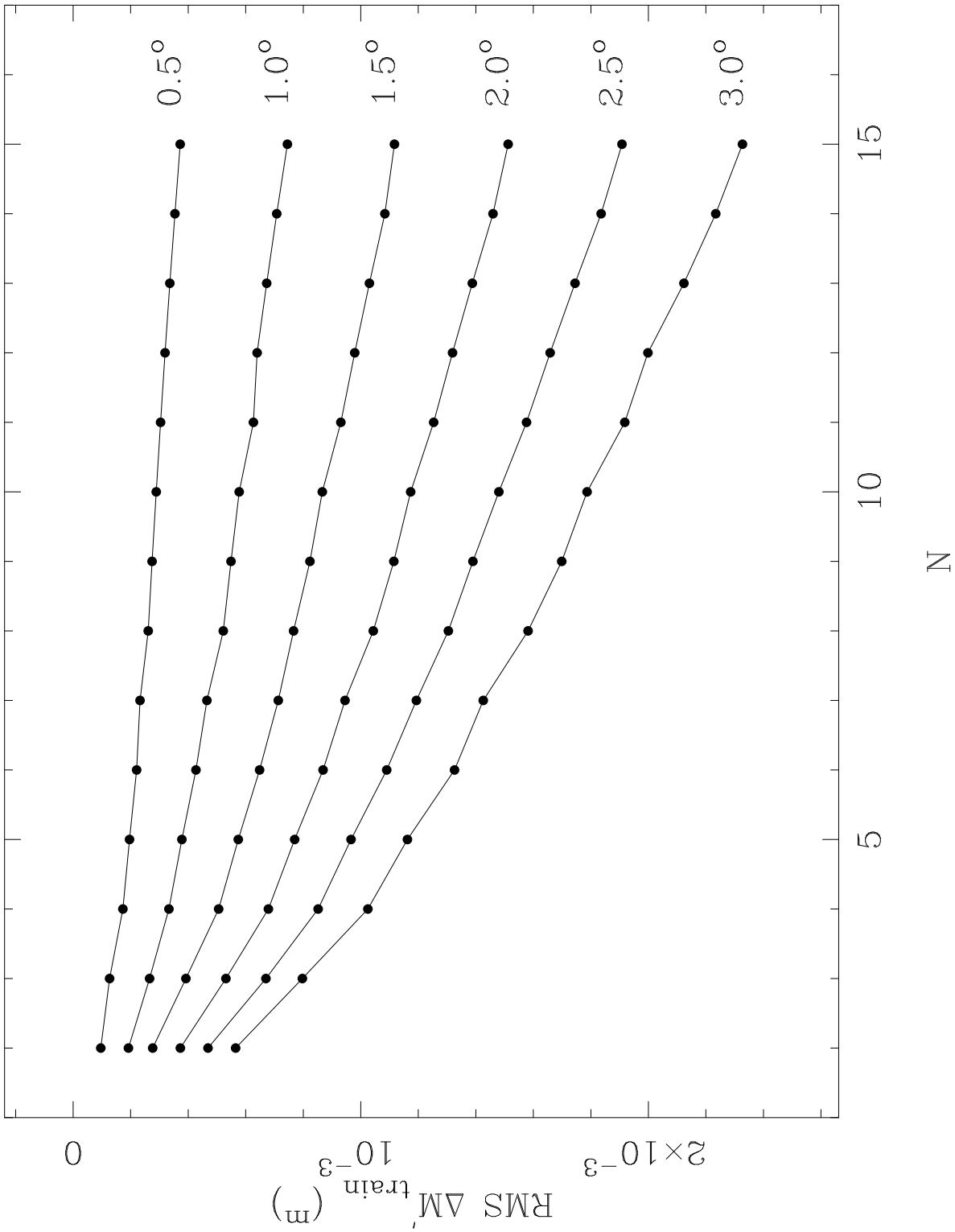

Fig. 7.— RMS $\Delta M'_{train}$ versus N and σ_{angle}. No distribution bars are shown, to avoid confusion.

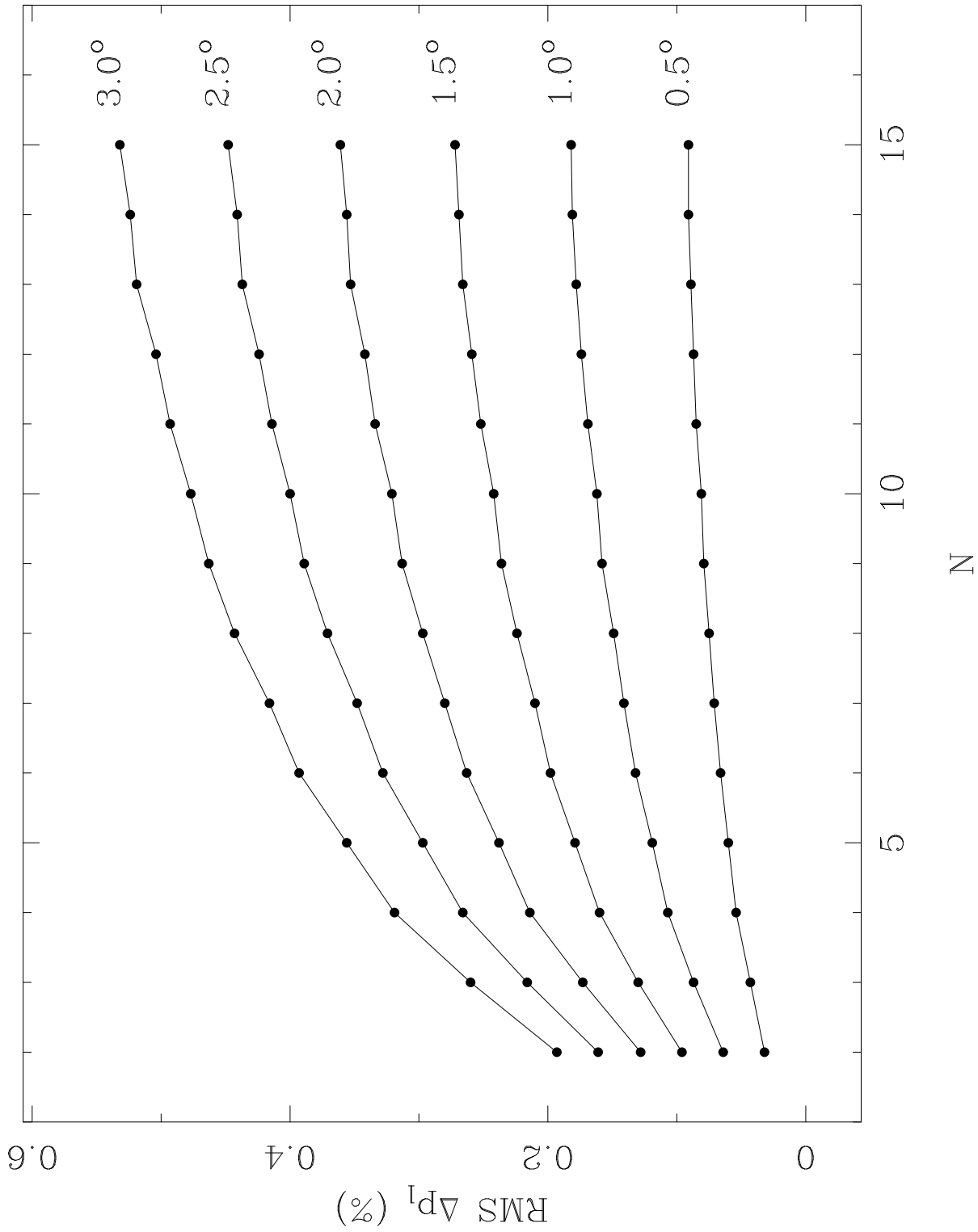

Fig. 8.— RMS Δp_l versus N and σ_{angle}. No distribution bars are shown, to avoid confusion.

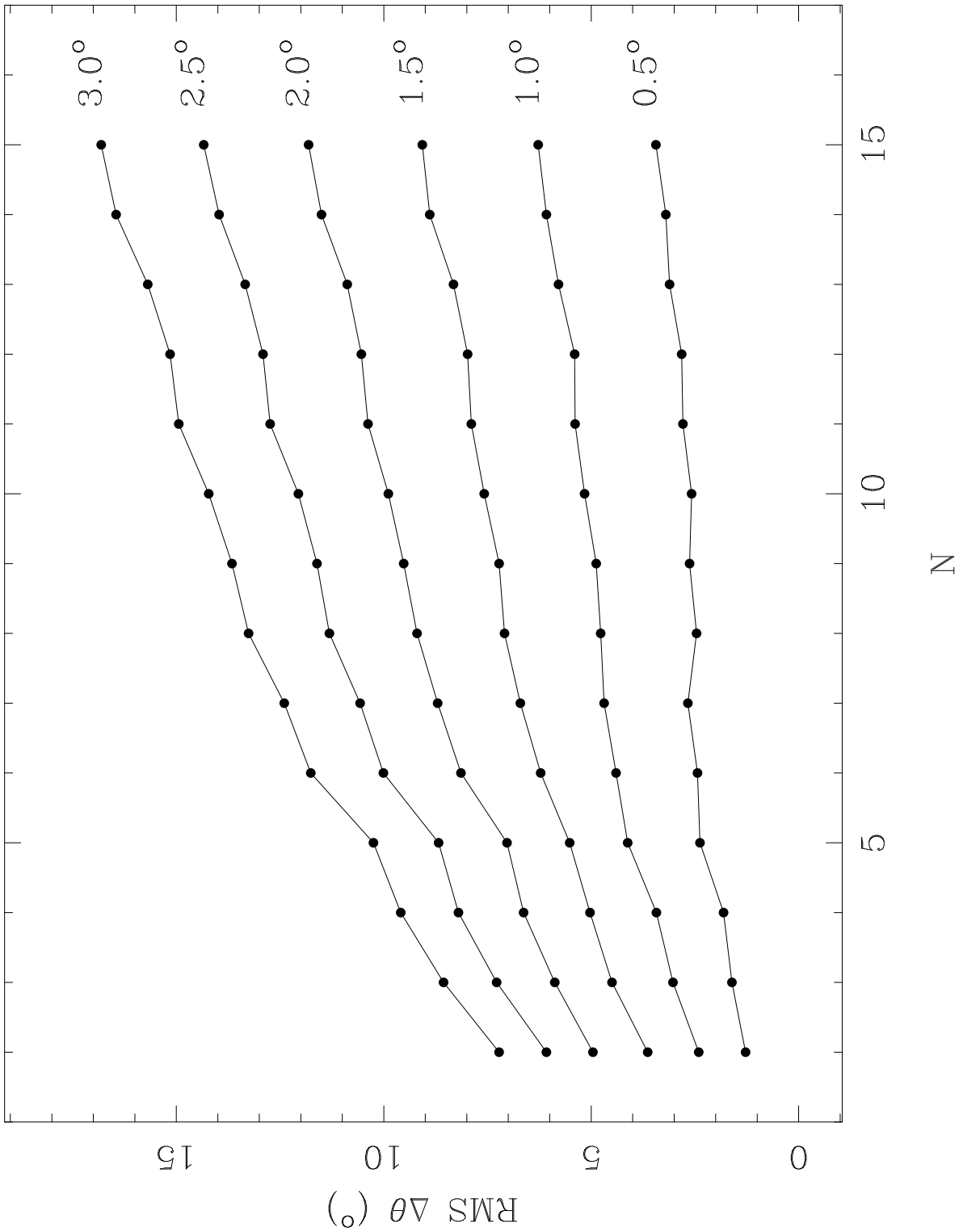

Fig. 9.— RMS $\Delta\theta$ versus N and σ_{angle}. No distribution bars are shown, to avoid confusion.

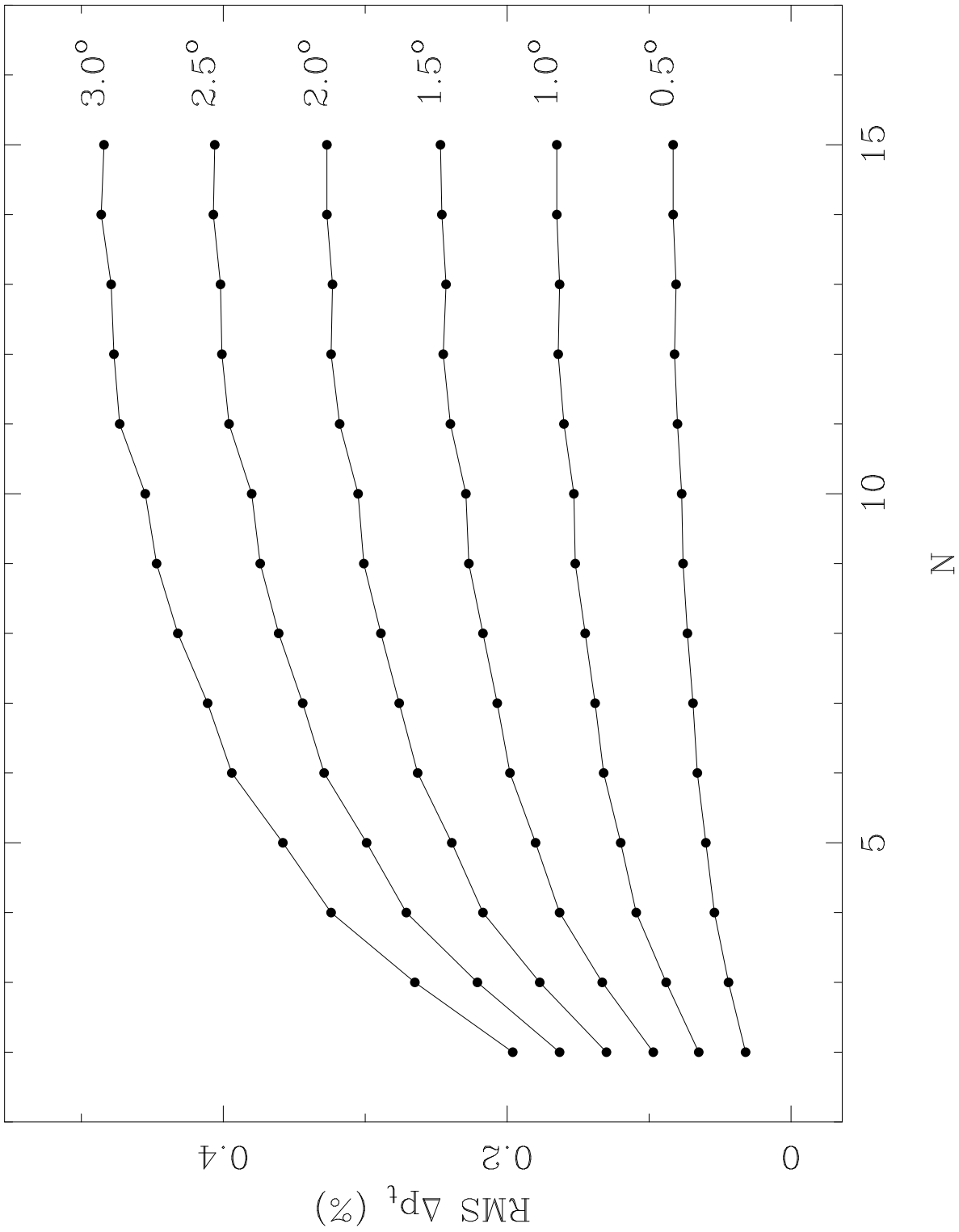

Fig. 10.— RMS Δp_t versus N and σ_{angle}. No distribution bars are shown, to avoid confusion.

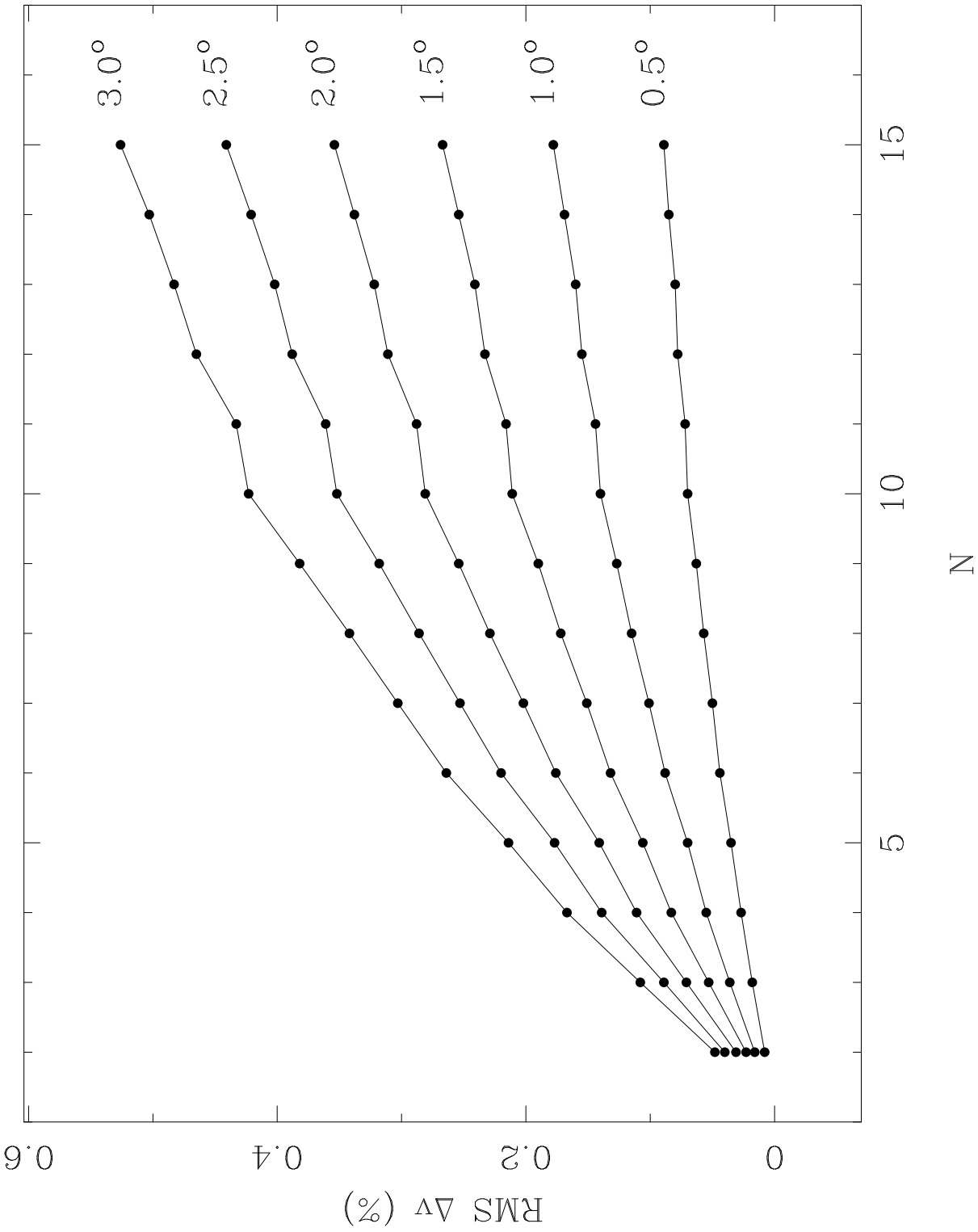

Fig. 11.— RMS Δv versus N and σ_{angle}. No distribution bars are shown, to avoid confusion.

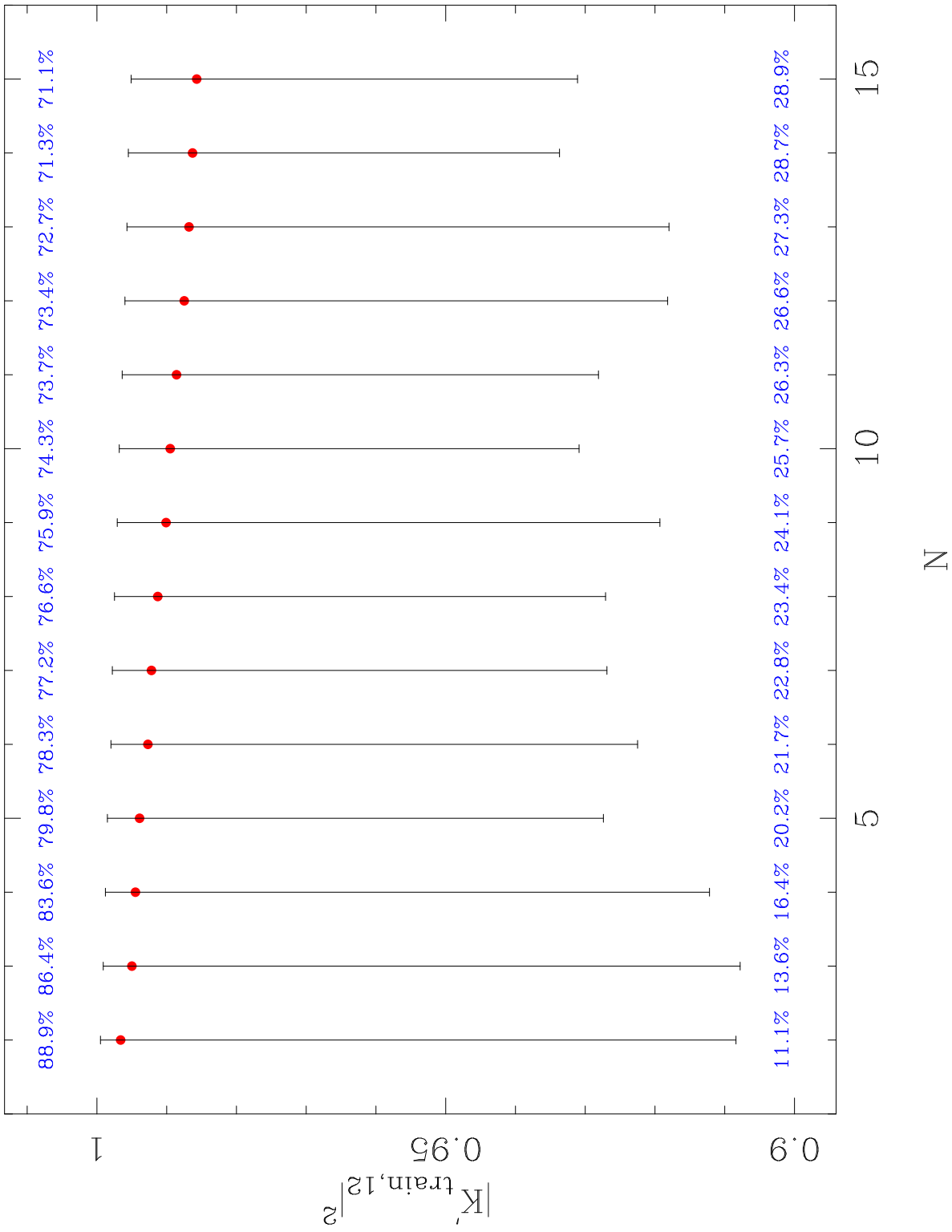

Fig. 12.— $K'_{train,12}$ versus N, $\sigma_{angle} = 1.0°$. Distribution bars, not errors of the means, are shown. The percentages represent the probabilities above and below the mean.

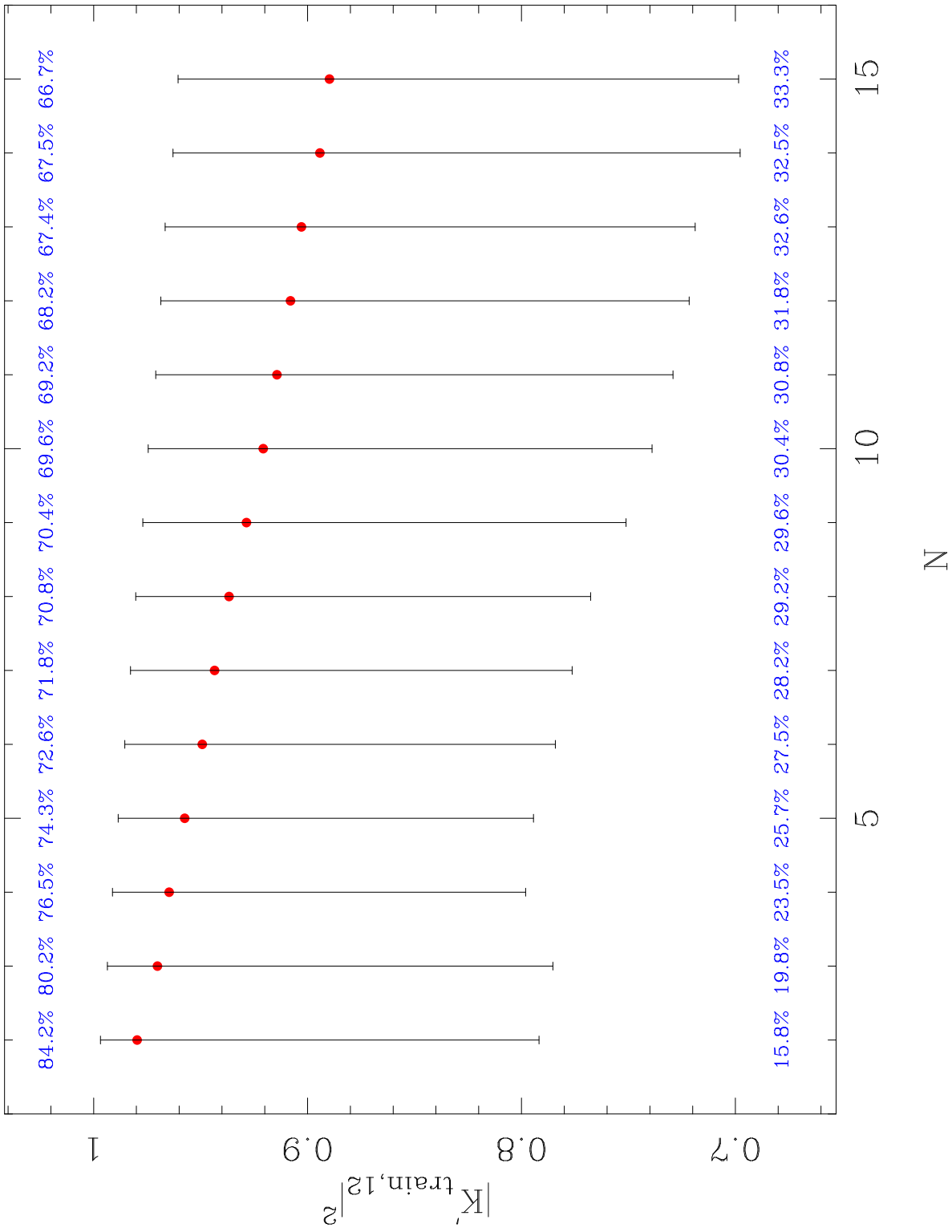

Fig. 13.— $K'_{train,12}$ versus N, $\sigma_{angle} = 3.0°$. Distribution bars, not errors of the means, are shown. The percentages represent the probabilities above and below the mean.

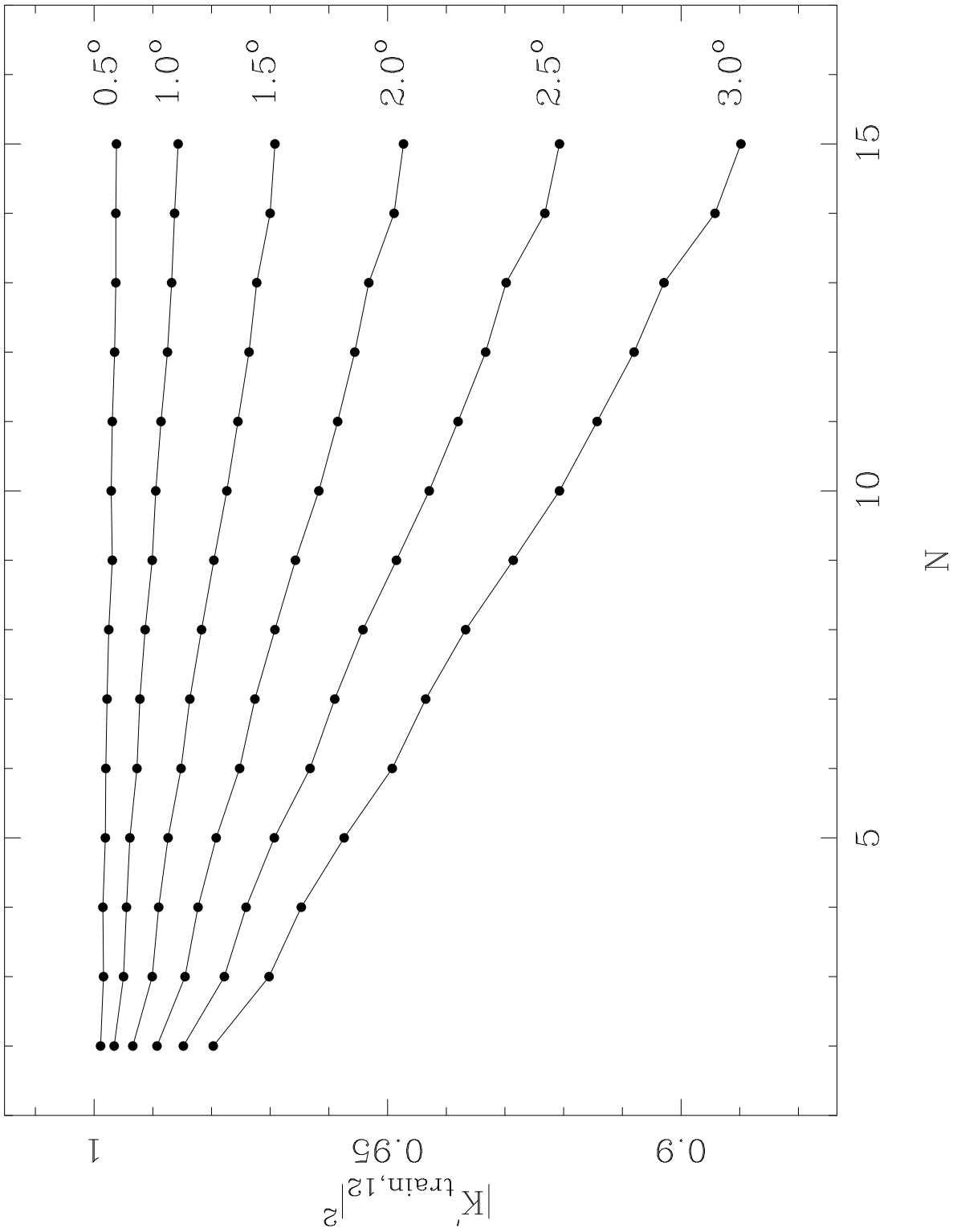

Fig. 14.— $K'_{train,12}$ versus N and σ_{angle}. No distribution bars are shown, to avoid confusion.

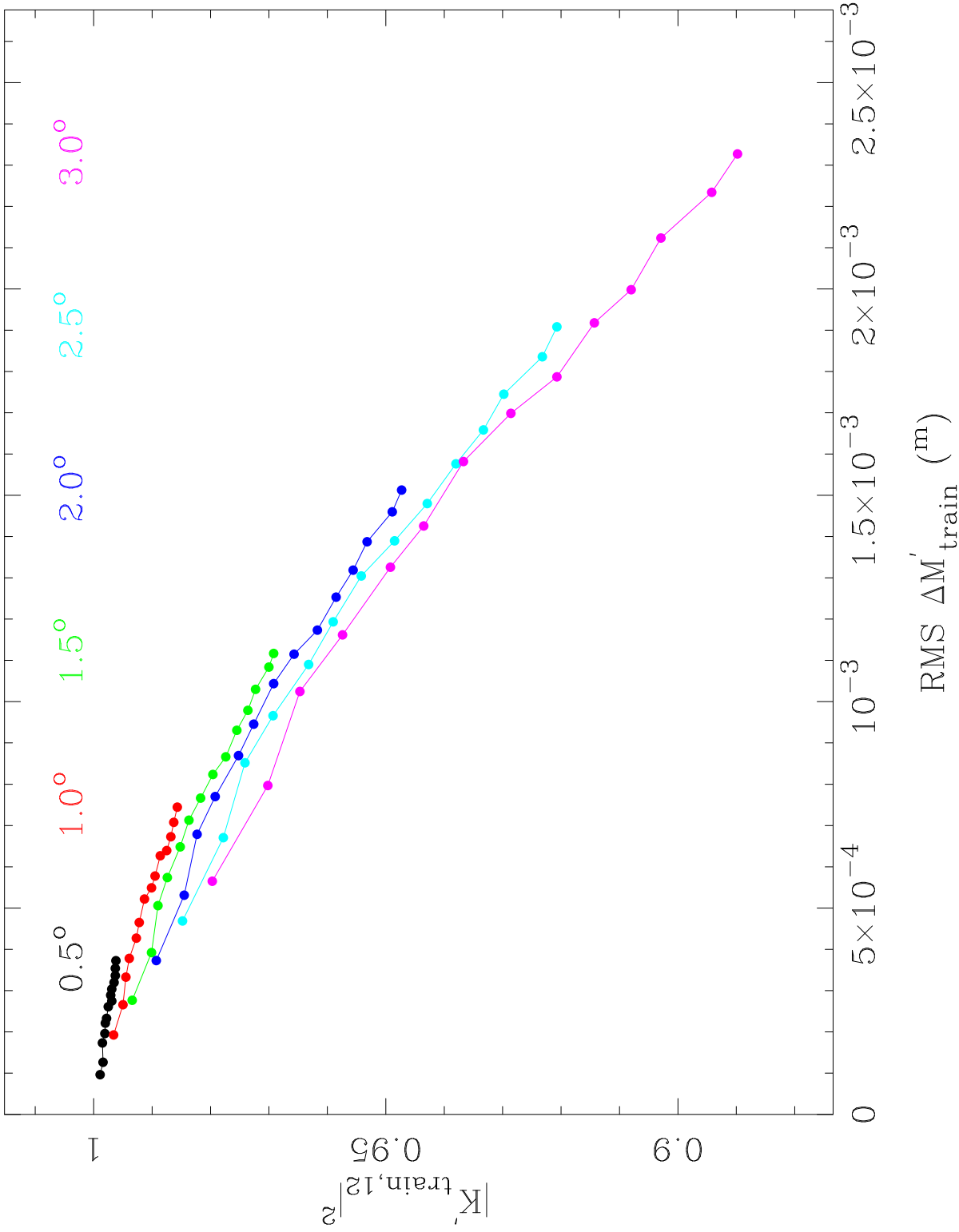

Fig. 15.— $K'_{train,12}$ versus RMS $\Delta M'_{train}$. For each σ_{angle} curve, the number of mirrors increases toward the right. No distribution bars are shown, to avoid confusion.

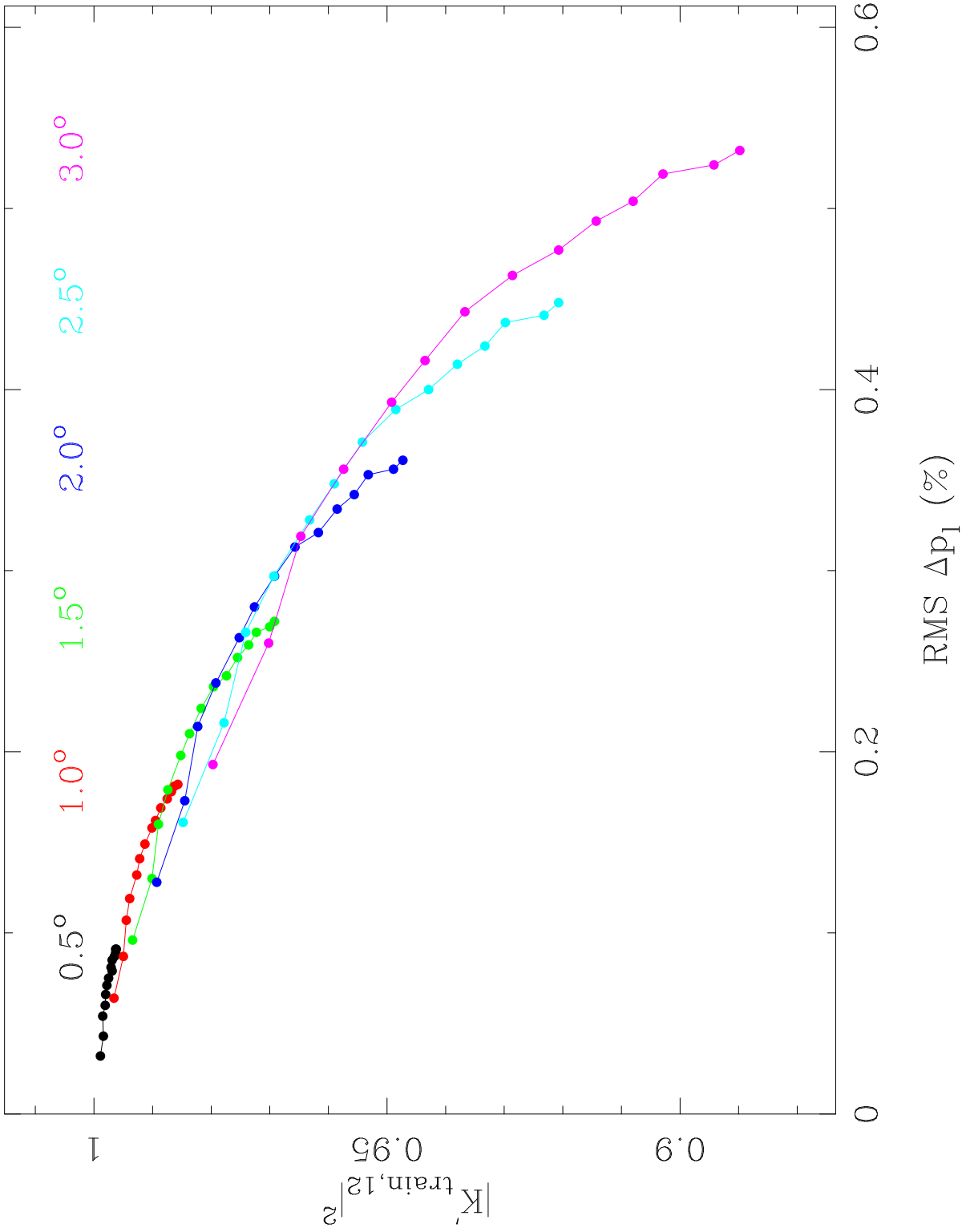

Fig. 16.— $K'_{train,12}$ versus RMS Δp_l. For each σ_{angle} curve, the number of mirrors increases toward the right. No distribution bars are shown, to avoid confusion.

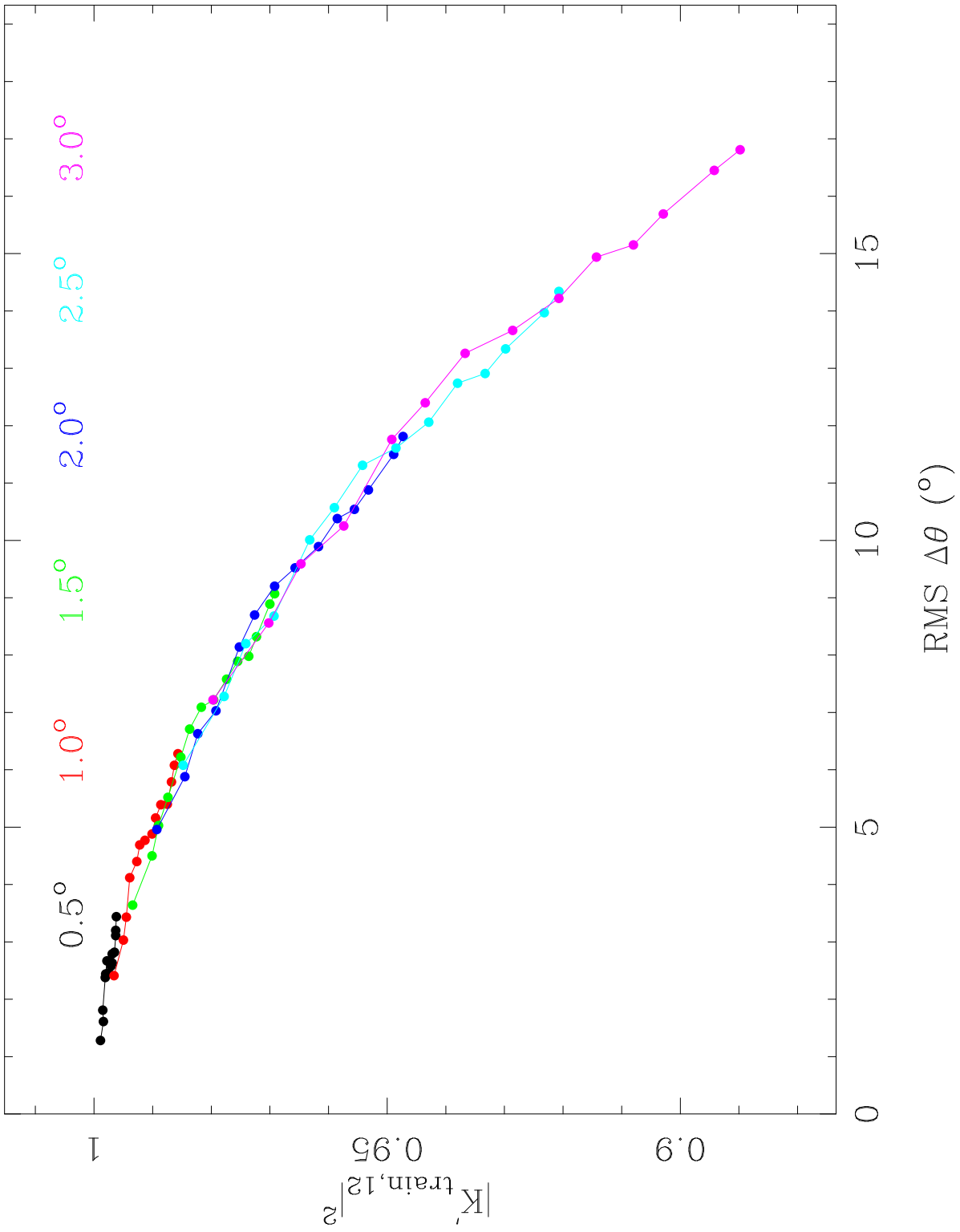

Fig. 17.— $K'_{train,12}$ versus RMS $\Delta\theta$. For each σ_{angle} curve, the number of mirrors increases toward the right. No distribution bars are shown, to avoid confusion.

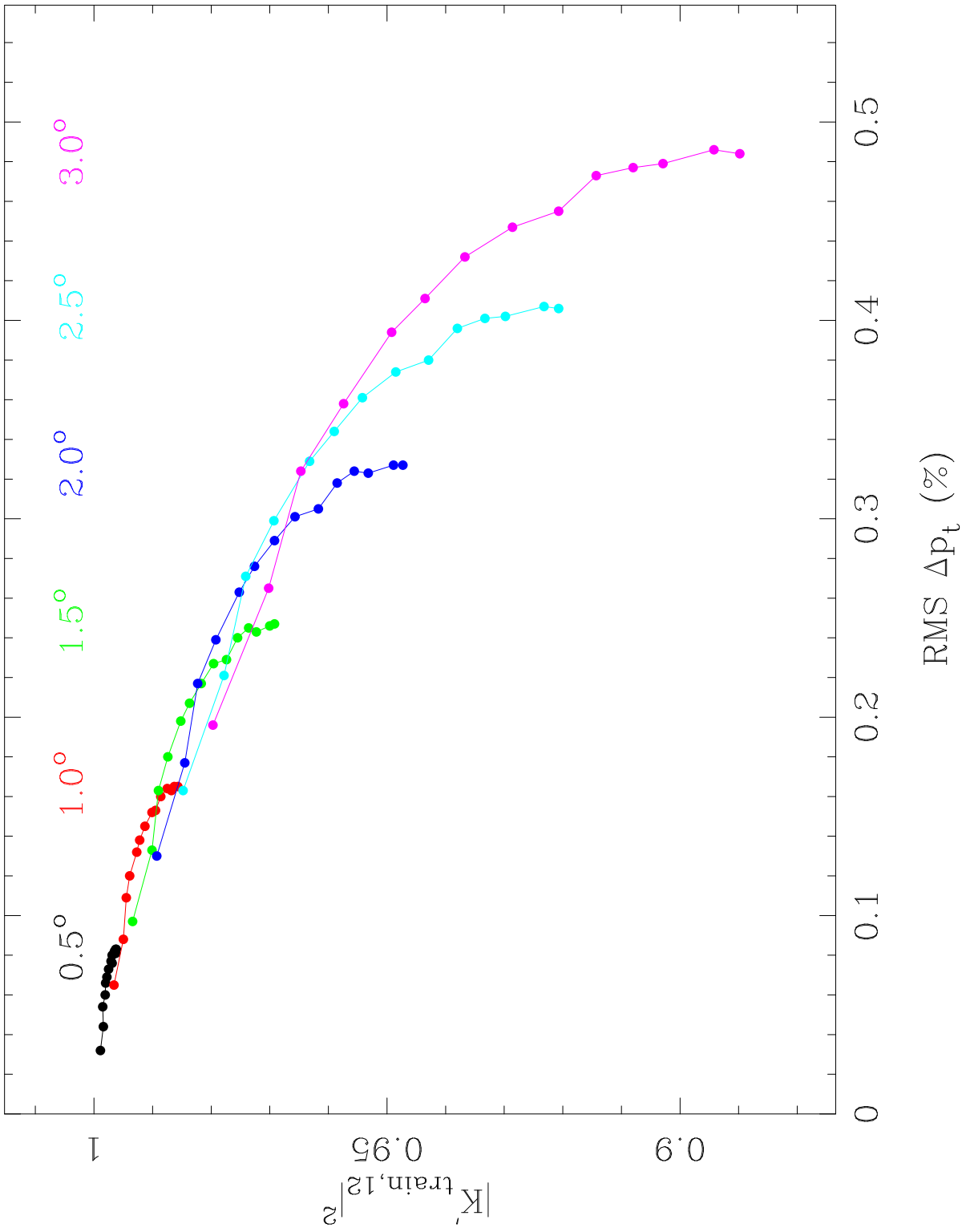

Fig. 18.— $K'_{train,12}$ versus RMS Δp_t. For each σ_{angle} curve, the number of mirrors increases toward the right. No distribution bars are shown, to avoid confusion.

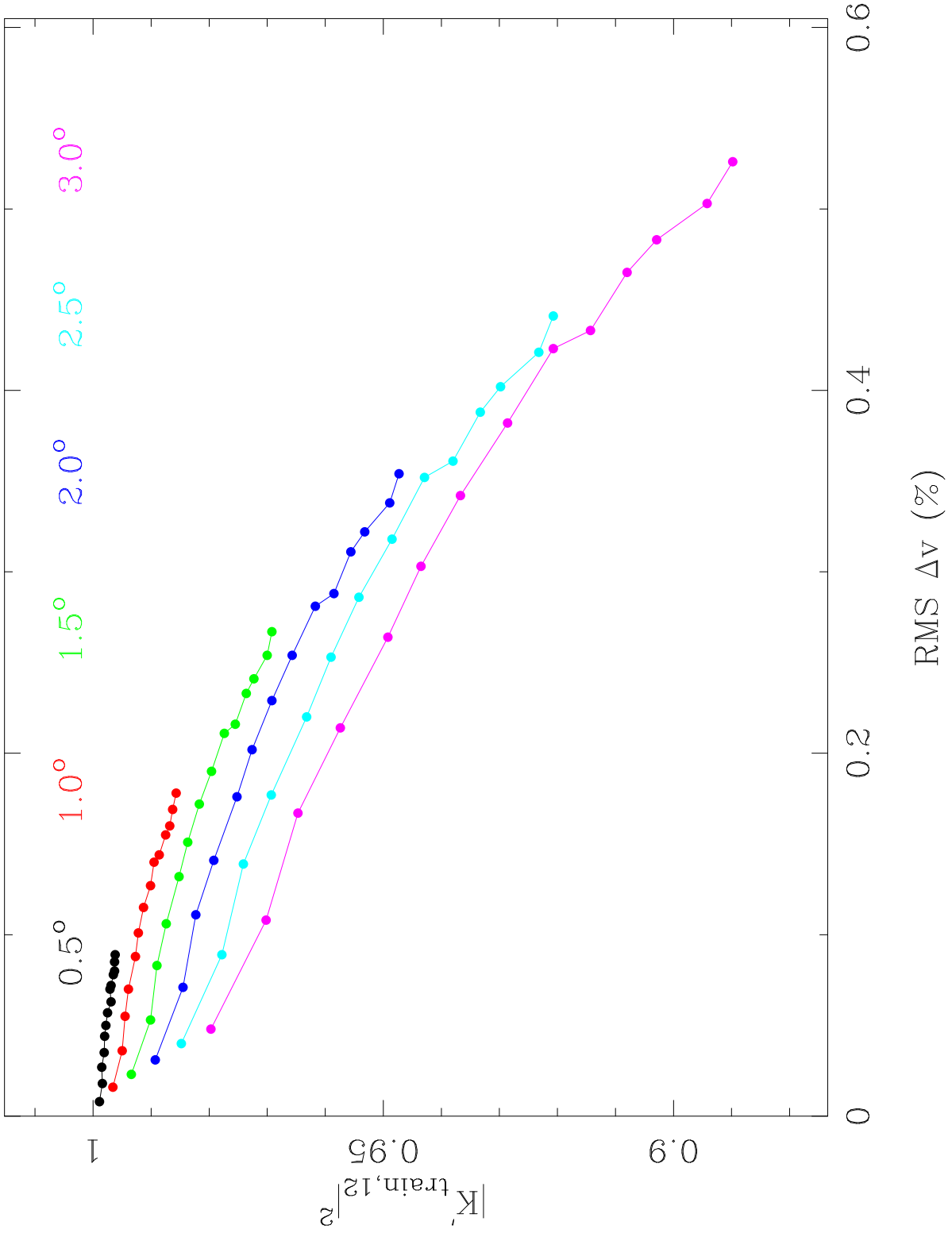

Fig. 19.— $K'_{train,12}$ versus RMS Δv. For each σ_{angle} curve, the number of mirrors increases toward the right. No distribution bars are shown, to avoid confusion.